2

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

THE USE OF EXPONENTIAL SMOOTHING TO PRODUCE YEARLY UPDATES OF LOSS RATE ESTIMATES IN MARINE CORPS MANPOWER MODELS

by

Daniel L. Hogan, Jr.

June 1986

Thesis Advisor: Robert R. Read

Approved for public release; distribution is unlimited.

SECURITY CLASSIFICATION OF THIS PAGE   AD-A171774

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION: UNCLASSIFIED

1b. RESTRICTIVE MARKINGS:

2a SECURITY CLASSIFICATION AUTHORITY:

2b DECLASSIFICATION / DOWNGRADING SCHEDULE:

3 DISTRIBUTION / AVAILABILITY OF REPORT: Approved for public release; distribution is unlimited.

4 PERFORMING ORGANIZATION REPORT NUMBER(S):

5 MONITORING ORGANIZATION REPORT NUMBER(S):

6a. NAME OF PERFORMING ORGANIZATION: Naval Postgraduate School

6b OFFICE SYMBOL (If applicable): Code 55

7a. NAME OF MONITORING ORGANIZATION: Naval Postgraduate School

6c. ADDRESS (City, State, and ZIP Code): Monterey, California 93943-5000

7b. ADDRESS (City, State, and ZIP Code): Monterey, California 93943-5000

8a NAME OF FUNDING / SPONSORING ORGANIZATION:

8b. OFFICE SYMBOL (If applicable):

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER:

8c ADDRESS (City, State, and ZIP Code):

10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO

11 TITLE (Include Security Classification): THE USE OF EXPONENTIAL SMOOTHING TO PRODUCE YEARLY UPDATES OF LOSS RATE ESTIMATES IN MARINE CORPS MANPOWER MODELS

12 PERSONAL AUTHOR(S): Hogan, Daniel L.

13a TYPE OF REPORT: Master's Thesis

13b TIME COVERED: FROM ______ TO ______

14 DATE OF REPORT (Year, Month, Day): 1986 June

15 PAGE COUNT: 105

16 SUPPLEMENTARY NOTATION:

17 COSATI CODES

FIELD	GROUP	SUB-GROUP

18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Exponential Smoothing, Attrition Rates, James-Stein Estimators, Limited Translation James-Stein Estimators, Transformed Scale Cell Average

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

The use of exponential smoothing to perform yearly updating of attrition rates is examined and has merit. It shows enormous flexibility in adjusting to changes in the environment affecting the attrition rates, and it displays almost as much accuracy as the method it is intended to replace while using thousands of times less data.

A secondary purpose of this study is fulfilled in conforming that the current aggregate methods are outperformed by maximum likelihood estimation, transform cell James-Stein. None of these four methods is dominant overall, but all are improvements over the estimation system now employed.

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS

21 ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED

22a NAME OF RESPONSIBLE INDIVIDUAL: Robert R. Read

22b TELEPHONE (Include Area Code): (408) 646-2382

22c OFFICE SYMBOL: Code 55Re

DD FORM 1473, 84 MAR   83 APR edition may be used until exhausted. All other editions are obsolete.   SECURITY CLASSIFICATION OF THIS PAGE

Approved for public release; distribution is unlimited.

The Use of Exponential Smoothing To Produce Yearly Updates of Loss Rate Estimates in Marine Corps Manpower Models

by

Daniel L. Hogan, Jr.
First Lieutenant, United States Army
B.S., United States Military Academy, 1984

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 1986

Author: Daniel L. Hogan, Jr.
Daniel L. Hogan, Jr.

Approved by: R. R. Read
Robert R. Read, Thesis Advisor

Jack B. Gafford
Jack B. Gafford, Second Reader

Alan R. Washburn, Chairman,
Department of Operations Research

Kneale T. Marshall
Kneale T. Marshall
Dean of Information and Policy Sciences

ABSTRACT

The use of exponential smoothing to perform yearly updating of attrition rates is examined and has merit. It shows enormous flexibility in adjusting to changes in the environment affecting the attrition rates, and it displays almost as much accuracy as the method it is intended to replace while using thousands of times less data.

A secondary purpose of this study is fulfilled in confirming that the current aggregate methods are outperformed by maximum likelihood estimation, transform cell scale average, James-Stein, and limited translation James-Stein. None of these four methods is dominant overall, but all are improvements over the estimation system now employed.

TABLE OF CONTENTS

I. INTRODUCTION 10

A. PURPOSE 10

B. BACKGROUND 10

C. PREVIOUS WORK 11

D. AGGREGATION METHOD 12

E. ESTIMATION METHODS 13

F. RESULTS 17

II. YEARLY UPDATING METHODS 18

A. GENERAL 18

B. CRITERIA FOR MODEL SELECTION 18

1. Accuracy 19

2. Size of Data 19

3. Flexibility to the Rate of Response 20

C. SELECTION OF THE YEARLY UPDATING MODEL 20

1. Candidate One -- Tucker/Robinson's Methods 20

2. Candidate Two -- Moving Average 21

3. Candidate Three -- Exponential Smoothing 22

III. EXPONENTIAL SMOOTHING 26

A. GENERAL 26

B. FIGURES OF MERIT 28

C. FINDING THE SMOOTHING CONSTANTS 30

D. LENGTH OF BASE PERIOD 31

1. Combat Support 32

2. Ground Combat 33

3. Aviation 34

E. STABILITY OF THE SMOOTHING CONSTANTS 36

1. Within Base Period Lengths 36

2. Between Base Period Lengths 50

IV. RESULTS 65

A. GENERAL 65

B. COMPARISON WITH RESULTS FROM EARLIER WORKS .. 65

1. Aviation 66

2. Combat Support 69

3. Ground Combat 72

C. ATTRITION RATES 73

V. CONCLUSIONS AND RECOMMENDATIONS 88

A. CONCLUSIONS 88

B. RECOMMENDATIONS 89

APPENDIX A: FUNCTIONS USED IN CALCULATIONS 91

APPENDIX B: FREEMAN-TUKEY ARCSINE TRANSFORMATION 98

1. GENERAL 98

2. THE TRANSFORMATION 98

3. THE INVERSE TRANSFORMATION 99

APPENDIX C: ANALYSIS OF OPTIMAL ALPHA OF TRANSFORMED AND ORIGINAL SCALES 100

1. GENERAL 100

2. AVIATION 101

3. COMBAT SUPPORT 101

4. GROUND COMBAT 102

LIST OF REFERENCES 103

INITIAL DISTRIBUTION LIST 104

LIST OF TABLES

1. DETERMINATION OF OPTIMAL BASE PERIOD -- COMBAT SUPPORT 32
2. DETERMINATION OF OPTIMAL BASE PERIOD -- GROUND COMBAT 33
3. DETERMINATION OF OPTIMAL BASE PERIOD -- AVIATION 34
4. EXPONENTIAL SMOOTHING AVIATION FIGURES OF MERIT 66
5. EXPONENTIAL SMOOTHING COMBAT SUPPORT FIGURES OF MERIT 67
6. EXPONENTIAL SMOOTHING GROUND COMBAT FIGURES OF MERIT 68
7. FOUR-YEAR BASE ESTIMATE AVIATION FIGURES OF MERIT 69
8. FOUR-YEAR BASE ESTIMATE COMBAT SUPPORT FIGURES OF MERIT 70
9. FOUR-YEAR BASE ESTIMATE GROUND COMBAT FIGURES OF MERIT 71
10. AVIATION ATTRITION RATES FOR 1ST LTS 74
11. AVIATION ATTRITION RATES FOR LTCOLS 75
12. COMBAT SUPPORT ATTRITION RATES FOR 1ST LTS CODE 07 ENGINEERS 76
13. COMBAT SUPPORT ATTRITION RATES FOR 1ST LTS CODE 13 COMMUNICATIONS 77
14. COMBAT SUPPORT ATTRITION RATES FOR 1ST LTS CODE 20 MOTOR TRANSPORT 78
15. COMBAT SUPPORT ATTRITION RATES FOR LTCOLS CODE 07 ENGINEERS 79
16. COMBAT SUPPORT ATTRITION RATES FOR LTCOLS CODE 13 COMMUNICATIONS 80
17. COMBAT SUPPORT ATTRITION RATES FOR LTCOLS CODE 20 MOTOR TRANSPORT 81
18. GROUND COMBAT ATTRITION RATES FOR 1ST LTS CODE 03 INFANTRY 82
19. GROUND COMBAT ATTRITION RATES FOR 1ST LTS CODE 05 ARTILLERY 83
20. GROUND COMBAT ATTRITION RATES FOR 1ST LTS CODE 10 TANKS 84

21. GROUND COMBAT ATTRITION RATES FOR LTCOLS CODE 03 INFANTRY 85

22. GROUND COMBAT ATTRITION RATES FOR LTCOLS CODE 05 ARTILLERY 86

23. GROUND COMBAT ATTRITION RATES FOR LTCOLS CODE 10 TANKS 87

24. COMPARISON OF TRANSFORMED AND ORIGINAL ALPHA AVIATION AGGREGATE 100

25. COMPARISON OF TRANSFORMED AND ORIGINAL ALPHA COMBAT SUPPORT AGGREGATE 101

26. COMPARISON OF TRANSFORMED AND ORIGINAL ALPHA GROUND COMBAT AGGREGATE 102

LIST OF FIGURES

1.1 James-Stein Shrinkage 15
1.2 Limited Translation James-Stein Shrinkage 16
3.1 Within 3-Year Base Stability -- Aviation 1LTS . . . 37
3.2 Within 3-Year Base Stability -- Aviation LTCOLS . . 38
3.3 Within 3-Year Base Stability -- Cbt. Spt. 1LTS . . . 39
3.4 Within 3-Year Base Stability -- Cbt. Spt. LTCOLS 40
3.5 Within 3-Year Base Stability -- Grd. Cbt. 1LTS . . . 41
3.6 Within 3-Year Base Stability -- Grd. Cbt. LTCOLS 42
3.7 Within 4-Year Base Stability -- Aviation 1LTS . . . 43
3.8 Within 4-Year Base Stability -- Aviation LTCOLS . . 44
3.9 Within 4-Year Base Stability -- Cbt. Spt. 1LTS . . . 45
3.10 Within 4-Year Base Stability -- Cbt. Spt. LTCOLS 46
3.11 Within 4-Year Base Stability -- Grd. Cbt. 1LTS . . . 47
3.12 Within 4-Year Base Stability -- Grd. Cbt. LTCOLS 48
3.13 Between VY One Stability -- Aviation 1LTS 51
3.14 Between VY Two Stability -- Aviation 1LTS 52
3.15 Between VY One Stability -- Aviation LTCOLS 53
3.16 Between VY Two Stability -- Aviation LTCOLS 54
3.17 Between VY One Stability -- Combat Spt. 1LTS 55
3.18 Between VY Two Stability -- Combat Spt. 1LTS 56
3.19 Between VY One Stability -- Combat Spt. LTCOLS . . . 57
3.20 Between VY Two Stability -- Combat Spt. LTCOLS . . . 58
3.21 Between VY One Stability -- Ground Combat 1LTS . . . 59
3.22 Between VY Two Stability -- Ground Combat 1LTS . . . 60
3.23 Between VY One Stability -- Ground Combat LTCOLS 61
3.24 Between VY Two Stability -- Ground Combat LTCOLS 62

A.1 APL Function ALPHAHAT 92
A.2 APL Function A5B 93
A.3 APL Function ESTIM 94
A.4 APL Function BINPREP 94
A.5 APL Function SUMSQ 95
A.6 APL Function MLE 95
A.7 APL Function XFOUR 95
A.8 APL Function RISKO 96
A.9 APL Function RISKT 97
A.10 APL Function BINCONV 97

I. INTRODUCTION

A. PURPOSE

This study continues the work started by Major D.D. Tucker [Ref. 1] and continued by Major John R. Robinson [Ref. 2] in their respective theses submitted at the Naval Postgraduate School (NPS) in September 1985 and March 1986. Their works [Refs. 1,2] dealt with obtaining better attrition rate estimates for the Marine Corps officer manpower model than the ones currently in use. Both Tucker and Robinson in the "Recommendations" section of their theses [Ref. 1: p. 72] [Ref. 2: p. 69] stated that further work on how to update the estimates from year to year was needed. This study investigates this "yearly updating" problem.

The primary purpose of this work is to investigate the efficacy of the exponential smoothing model as a yearly updating model for Marine Corps attrition rates. The reasons for choosing the exponential smoothing model for study are outlined in Chapter II.

A secondary objective of this study will be to compare the performance of the attrition rate estimators introduced by Robinson [Ref. 2], which are introduced later in this chapter, to the current Marine Corps' estimator.

B. BACKGROUND

Much of the background detail for this study is well-documented in previous works. Tucker [Ref. 1: pp. 15-38, 128-138] explains the Marine Corps Officer attrition and promotion structure, the structure of the data base used by the Marine Corps in its officer planning process, the process itself, and how the data from the Navy Personnel Research and Development Center was transferred to the NPS computer system. Robinson's study [Ref. 2: pp. 11,14-20]

contains a complete summary of the estimation methods and aggregation procedure he used. For continuity, these estimation methods and this aggregation scheme are used in the present work.

Robinson's thesis [Ref. 2: pp. 74-75] also includes an explanation of the Freeman-Tukey arcsine transformation, which is used to stabilize the variance of the empirical loss rate estimates. The empirical process is assumed to have a binomial(n,p) distribution with parameters

n = central inventory for the year

p = probability that an inventory unit leaves the Marine Corps during the year.

The Freeman-Tukey arcsine transformation provides a second scale for comparison of the estimators within which the estimates have a more stable variance, and additionally, behave more like a normal distribution (see Appendix B).

Both Tucker and Robinson [Refs. 1,2] used data from years 1977-1980 to obtain their loss rate estimates, and then used years 1981-1982 to validate them. Robinson [Ref. 2] also had year 1983 to use for validation of his estimates. Their estimates [Refs. 1,2] worked fairly well in predicting attrition one year into the future, but their success faded noticeably as the they tried to predict two or three years into the future. Thus there appears to be a time-varying component in the attrition rates. This leads to the problem of updating the loss rate estimates each year in order to better forecast future attrition.

C. PREVIOUS WORK

Major Tucker [Ref. 1] showed that the James-Stein shrinkage estimator was better than both the current method used by the Marine Corps and maximum likelihood estimators. He also showed that the James-Stein technique will provide estimates for those small cells which have no attrition,

i.e., those cells whose maximum likelihood estimator must be zero [Ref. 1]. Also, Tucker stated in his summary that the small inventory cells present a problem in loss rate estimation because "some of the cells are empty for structural reasons while others are empty by chance" [Ref. 1: p. 70].

Major Robinson [Ref. 2] tried to combat the "small cell" problem, i.e., the problem with estimating attrition rates for those cells with low inventory figures, by introducing the limited translation James-Stein technique of Efron and Morris [Refs. 3,4]. He showed that this technique improves upon the James-Stein estimates used by Tucker [Ref. 1] in estimating the rates for small cells [Ref. 2]. Robinson [Ref. 2] also introduced a transformed scale cell average (TSCA), an estimator corresponding to zero shrinkage in the James-Stein technique, which in many cases outperformed all other estimators.

D. AGGREGATION METHOD

The United States Marine Corps Officer Corps numbers approximately 20,000. Each officer below the rank of brigadier general is cross-classified into one of 40 military occupational fields (OF), 31 lengths of service years (LOS), corresponding to 0 to 30 years in the Marine Corps, and 10 grades, from warrant officer 1 to colonel, for a total of 12,400 categories.

Almost half (6149) of these categories, or cells, are "structural zeroes" in inventory. "Structural zeroes" occur due to Marine Corps policy concerning promotions and because certain combinations of OF, LOS, and grade never occur, e.g., there are no colonels with just 2 years in service. These cells exist only in theory, not in practice, and are therefore not included in the feasible region of cells.

A vast majority of the 6251 feasible cells are low inventory cells. Because of this, it is quite difficult to

obtain useful stable attrition rate estimates for these cells and it is wasteful to try to treat each cell individually. There is much communality of behavior among clusters of cells, and the grouping of cells into aggregates of like characteristics can ease the bookkeeping burden as well as provide the desired stability. Ideally, aggregation schemes can be found for which the aggregates behave in a predictable manner and for which meaningful conclusions may be drawn from statistical tests. However, current Marine Corps practice groups the cells according to organizational and operational considerations, producing aggregates that will not necessarily conform to any specific statistical behavior.

H. Amin Elseramegy used the CART algorithm to find aggregations with predictable statistical behavior with encouraging results in his thesis submitted at NPS in December, 1985 [Ref. 5]. His results cannot be regarded as definitive because of operational considerations (e.g., excessive computer running time required some pre-aggregation), but can serve to guide future work.

The current Marine Corps manpower model places all occupational fields into four categories: aviation (OF 72, 75), combat support (OF 13, 25, 35), ground combat (OF 03, 08, 18) and other (includes 32 occupational fields). This aggregation scheme is used by both Tucker and Robinson [Refs. 1,2] and will be used again in the present work. For continuity, the aviation category will include only OF 75 as it did in both of their works [Refs. 1,2].

E. ESTIMATION METHODS

Robinson in his thesis [Ref. 2] compared six loss rate estimators. These estimators are:

1) Original Aggregate (AGG ORIG) -- the current Marine Corps estimation method. The occupational fields are placed into the four categories mentioned in the preceding section. Past attrition rates from 1977 to the present are subjectively weighted and averaged for each aggregate. The grand mean of the weighted

attrition rates serves as the loss rate estimate for all cells (OF, LOS, grade) within the aggregate. Both Tucker and Robinson [Refs. 1,2] found methods superior to this one.

2) Transformed Aggregate (AGG TRANS) -- computed by transforming the empirical attrition rates using the Freeman-Tukey equation and then calculating the mean of the transformed values within each aggregate. This is followed by an inversion to the original scale. Again, this is a single number used for all cells within the aggregate.

3) Maximum Likelihood Estimator (MLE) -- Is calculated by summing all leavers (over time) in a cell and dividing by the total cell inventory (over time) for the estimation period. This estimator is the MLE for the binomial distribution described in the previous section. There are problems with using the MLE in this setting. While the estimate is indeed unbiased, it is unstable due to the abundance of small cells causing the possibility of a wide range of values. Also, this MLE assumes that each year represents an identical population, while the data shows that a cell's behavior can change drastically within a few years' time. A complete discussion of the problems with using the MLE can be found in [Ref. 2: pp. 17-18].

4) Transformed Scale Cell Average (TSCA) -- computed by transforming the cell inventory and loss data and calculating the time average for each cell over the estimation period. Inversion of the results provides attrition estimates in the original scale. Use of the Freeman-Tukey transform holds down the variability of this estimator, which was mentioned above as a shortcoming of MLE. This estimate performed surprisingly well in Robinson's analysis [Ref. 2].

5) James-Stein Estimator (JS) -- operates from the basic notion that by "shrinking" estimates toward the grand mean, the size of the sum of squared residual errors will be lessened. The shrinking is applied to the TSCA estimator. An optimal shrinkage factor is found for each aggregate, and the cell means are shrunk toward the grand mean by that amount. See Figure 1.1. The optimal shrinkage factors used here are those found by Robinson in his thesis [Ref. 2: pp. 34-36]. Notice that since the shrinkage is done in transformed space, the assumptions of normality and stable variance required by James-Stein are less compromised.

6) Limited Translation James-Stein (LTJS) -- intuitively, it does not seem quite right to shrink all of the values towards the grand mean. After all, extreme values do occur occasionally, and their effects should be represented in an analysis of the data. This estimator deals with this problem by limiting the translation of extreme values toward the grand mean. From Figure 1.2 one sees that there is an interval about the grand mean within which full James-Stein shrinkage occurs, while outside this interval, the shrinkage is diminished. There is a factor, d, which controls the width of this interval. Robinson [Ref. 2: p.37] found the optimal values for this factor, and they are used in the present work. For a detailed explanation of this estimator, see Robinson's thesis [Ref. 2: pp22-25].

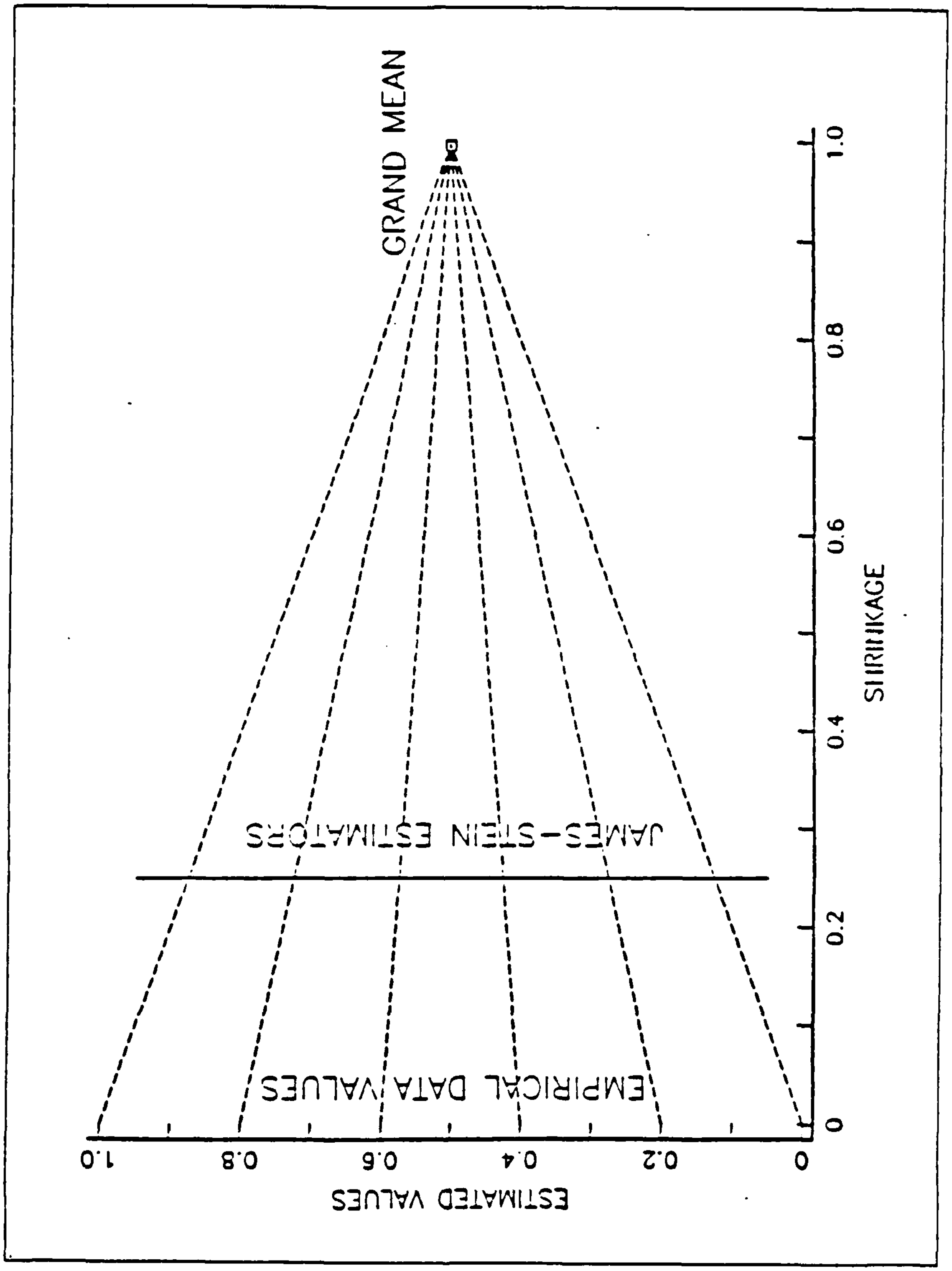

Figure 1.1 James-Stein Shrinkage

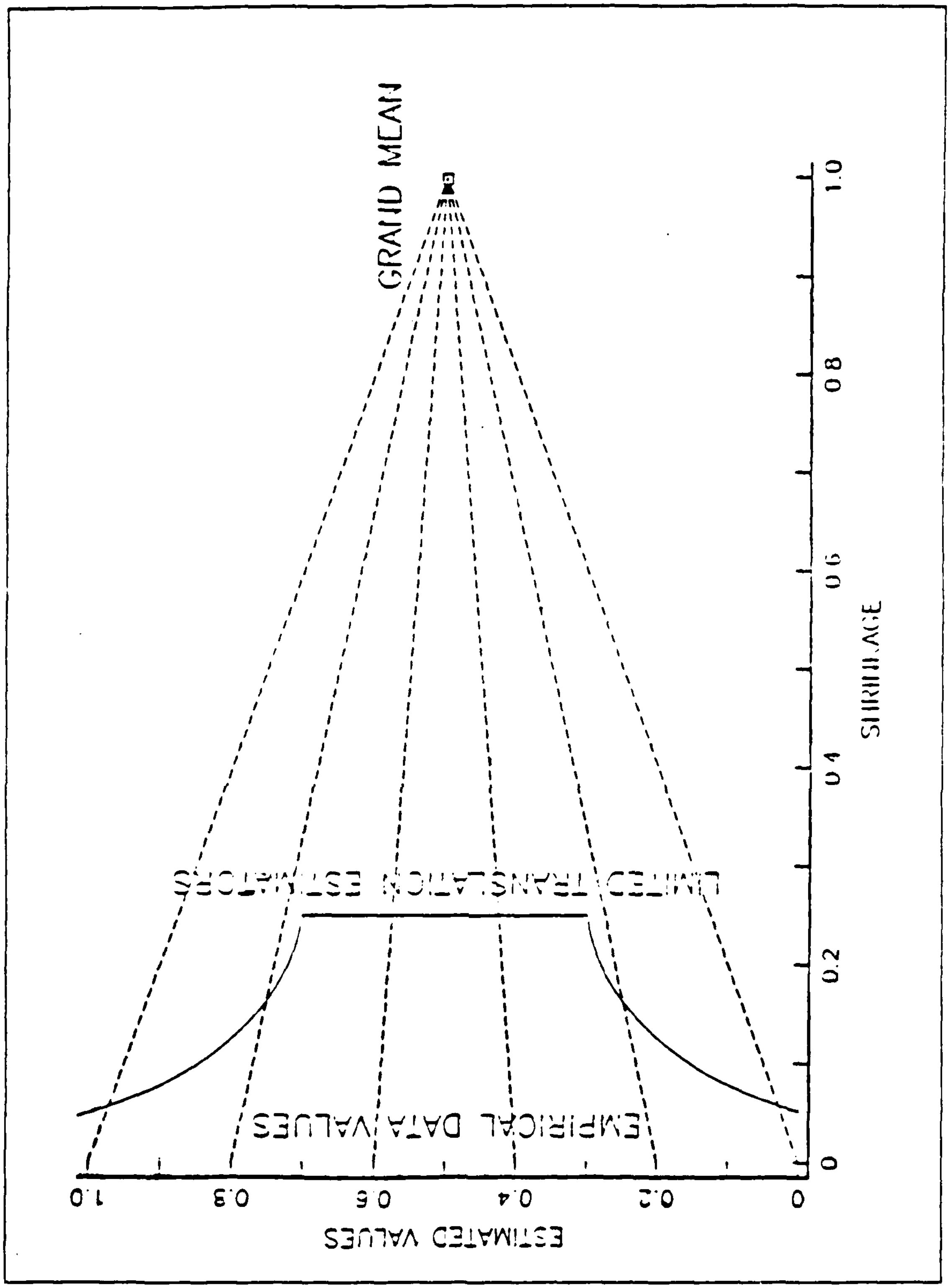

Figure 1.2 Limited Translation James-Stein Shrinkage

F. RESULTS

Exponential smoothing provides a valid yearly updating model for the Marine Corps attrition rates in all cases. The estimates produced by exponential smoothing are, more often than not, better than those produced by the methods of Robinson [Ref. 2], and require far less data. This model shows its extreme flexibility by pointing out an external change in the aviation environment that occurred in 1981 and actually outperforms the Robinson estimates for this aggregate.

In this study, as in Robinson's [Ref. 2], the TSCA, JS, and LTJS estimators forecast attrition rates better than the current method used by the Marine Corps. The maximum likelihood estimator performs well in the transformed scale, but has extremely large errors in original scale in some cases. This is due, in part, to the unsuitability in these cases of using the optimal smoothing constant values for transformed scale to produce smoothed estimates in original scale for this estimate (see Appendix C). However, as in Robinson's work [Ref. 2], no estimation method emerges as the clear-cut "best" choice.

The ability of exponential smoothing to update attrition rates is encouraging. However, having studied such a limited sample of data, we believe that its implementation should be delayed until further studies can be done when more data becomes available.

II. YEARLY UPDATING METHODS

A. GENERAL

A yearly updating model is one which allows newly received data to be combined with data from the past to update the estimate produced by that past data. By updating the estimate, forecasts produced by the prior estimate should be improved upon.

B. CRITERIA FOR MODEL SELECTION

There are many methods available to update parameter estimates, such as attrition rates, as more data becomes available. To choose among these methods, Robert Goodell Brown [Ref. 6] suggests using the following criteria: accuracy, simplicity of computation, and flexibility of rate of response. Of the three, only simplicity of computation is not a major concern for the yearly updating model. In 1963, when Brown published his book, simplicity of computation was important for a model because of the relative inefficiency of computers of that time as compared with those of today. Since today's computers are millions of times faster, we will not be concerned with this criterion in our model selection. However, a criterion which is important in choosing a yearly updating model is the amount of data that needs to be stored in order to produce the estimates.

Naturally there are other criteria which may be used to select a parameter estimation model. However, the three that seem most appropriate in determining the yearly updating model used in this study are:

1) accuracy
2) size of data
3) flexibility of rate of response .

1. Accuracy

Fulfilling the Marine Corps' need for more accurate attrition rate estimates is one of the primary purposes of this pilot study. More accurate forecasts are produced by such estimates, thereby avoiding some of the costly overages and underages in inventory which result from the use of the current estimation system.

One of the purposes of finding a yearly updating model is to improve upon the forecasts of Tucker and Robinson [Refs. 1,2], for which accuracy dropped off sharply after the first year. The yearly updating model selected should be able to forecast attrition rates one year in advance at least as well as Tucker and Robinson's [Refs. 1,2] did, and also improve greatly upon forecasting rates two or three years from the present.

In reality, it is the estimation of attrition rates two years from the time of the most recent data that should be of major concern. Loss and inventory data for a fiscal year is not generally available until halfway through the following fiscal year. Therefore, with the data available, the next time period in which attrition rates need to be forecasted is the year following the year in which the data is received, which is two years ahead of the most current available data. Thus the accuracy of the estimates for two years from the time of the most current data, for which Tucker and Robinson [Refs. 1,2] had little success, is one of the key figures with which the yearly updating model need be concerned.

2. Size of Data

This criterion is very important in choosing a yearly updating model. With 6251 feasible cells, it becomes unwieldy and costly in terms of computer storage space to log year after year of loss and inventory data. Thus, a model is sought which can use short summaries of the data and still produce valuable estimates.

3. Flexibility to the Rate of Response

In forecasting, when the current observation is different from the forecasted value, there are two possible explanations: random fluctuation; change in the pattern of the data. If the error is a random fluctuation in the data, then the forecasting technique should smooth out the fluctuation. In order for the model to do this, it should produce estimates based on a great deal of past data. However, if the error is due to a change in the pattern of the data, then past data is rendered irrelevant. The estimate should reflect only the recent processes.

Changes in Marine Corps policies in handling its officer corps, which occur from time to time, often cause corresponding changes in attrition patterns. Given information about such changes, one should be able to detect a new pattern in the loss and/or inventory data resulting from them. Thus, we want our yearly updating model to be able to smooth out random fluctuations in the data during times when the attrition process is stable over a period of years, yet still be able to respond rapidly to new conditions. The model must be able to easily adjust the number and relative value of past observations in producing the estimate, using a fairly long series of data for an unchanging process and only the most recent observations when a change in the process occurs.

C. SELECTION OF THE YEARLY UPDATING MODEL

1. Candidate One -- Tucker/Robinson's Methods

Both Tucker and Robinson [Refs. 1,2] showed that the alternate estimators they introduced outperform the aggregation methods currently employed by the Marine Corps. They used four equally weighted years of data, 1977-1980, to produce their estimates [Refs. 1,2]. The question that arises from their findings is why not use the functions that Tucker and Robinson wrote [Refs. 1,2] to compute estimates based on five, six, seven, etc., years of data?

The answer to this question lies in the criteria for model selection outlined in the previous section, of which the continuing use of Tucker and Robinson's methods over time meet none. The accuracy of Tucker and Robinson's estimates [Refs. 1,2] dropped off sharply after the first year in most cases, and the estimate of attrition two years off is very important, as mentioned earlier. The amount of data used to produce their four-year estimates was tremendous; storing all of that data plus that of additional years would be totally inefficient. Finally, using their methods with more data is not responsive to changes in attrition patterns. Old data, equally weighted with recent data and which may no longer be relevant, is used in this technique to produce the loss rate estimates.

Therefore, using Tucker and Robinson's methods [Refs. 1,2] year after year on all of the data available since 1977 is not a very good alternative. However, their methods are valuable for providing a base estimate for the exponential smoothing model to be discussed later.

2. <u>Candidate Two -- Moving Average</u>

A moving average is simply the sum of the most recent N observations divided by N. For example, imagine that a basketball player scores 18, 15, 25, 22, 20, and 14 points in his first six games of the season. Let the scores be denoted by x_i, $i = 1,2,...,6$. His scoring average is the sum of these six scores divided by six, or 19.0 points per game. Call this M_6. Suppose he scores 24 points in his seventh game. Then, if the practice of finding his scoring average over the past six games is continued, his six-game moving average, M_7, can be computed by adding the scores from the most recent six games and dividing by six. Another way would be to subtract 1/6 of the score he achieved six games ago, or game one, and adding 1/6 of his total for the most recent game, game seven. This produces the six-game moving average

$$M_7 = M_6 + (x_7 - x_1)/N$$
$$= 19.0 + (24-18)/6 = 20.0.$$

Using this procedure to find attrition rates would be straightforward; one would just find the actual rates for the past N years and average them. This method does away with some of the problems with data size encountered with the previous candidate. Whereas continued use of Tucker and Robinson's methods [Refs. 1,2] required the storage of all of the inventory and loss data from 1977 forward, this model will compute the rates each year for aggregates of cells, and only these rates, which number N times the number of aggregates, need be stored. This method may also prove to be more accurate than the continual use of Tucker and Robinson's methods [Refs. 1,2] since data from more than N years ago, which may bear little resemblance to current data, is eliminated. However, it is difficult to change the rate of response using a moving average. If a change in the underlying distribution of the data occurs, it will take N years for the moving average to fully reflect this change. One might suggest to keep N small so that it will respond more quickly to changes, but by doing this, the greater accuracy produced by larger data sets is sacrificed. Thus, N must be chosen so as to compromise between these conflicting objectives.

The moving average is an improvement over the previous candidate, but it requires more data and is not as flexible as the candidate which follows, the exponential smoothing model. However, an understanding of it is helpful, as the exponential smoothing model is itself a type of moving average.

3. <u>Candidate Three -- Exponential Smoothing</u>

It is a disadvantage of the moving average that it has to carry all of the rates needed to compute it, albeit is a great improvement over the amount needed for the

continued use of Tucker and Robinson's methods [Refs. 1,2]. Exponential smoothing cuts back even further on the amount of data needed. Let us see how by continuing the example of the previous subsection.

Suppose now that after recording the 24 pcints scored by the basketball player in game seven, it is discovered that all records of the previous six scores have been destroyed, but the moving average, $M_6 = 19.0$, still remains. If the value of x_1 was known, M_7 could be computed. The best estimate we have for x_1 is that is was equal to the average, $M_6 = 19.0$. Using this estimate for x_1, a new estimate of the six-game scoring average can be computed:

$$\begin{aligned} Mhat_7 &= M_6 + (x_7 - M_6)/6 \\ &= (1/6)x_7 + (5/6)M_6 = 19.833 \end{aligned}$$

$Mhat_7$ is an estimate of the moving average M_7, and is called the smoothed value of the average. $Mhat_i$ will hereafter be referred to as S_i, S standing for "smoothing."

If the equation used to find $Mhat_7 = S_7$ is used to find each succesive estimate, the definition of the smoothed function of the observations is [Ref. 6: p. 101]:

$$S_t(x) = \alpha\, x_t + (1-\alpha)S_{t-1}(x) \tag{2.1}$$

The smoothing constant, α, is similar, but not exactly equal to the fraction 1/N used to find the moving average [Ref. 6: p. 101]. The operation which updates an estimate by adding a fraction α of the difference between the current observation and the previous estimate to that previous estimate is called exponential smoothing [Ref. 6: p. 101].

Exponential smoothing discounts past data based upon the size of the α parameter. How it does so can be seen by substituting for the previous smoothed value the equation that produced it from an even earlier smoothed value [Ref. 6: p. 101]:

$$S_t(x) = \alpha\, x_t + (1-\alpha)(\alpha\, x_{t-1} + (1-\alpha)S_{t-2}(x)) \qquad (2.2)$$
$$= \alpha\, x_t + \alpha\,(1-\alpha)(x_{t-1}) + \alpha\,(1-\alpha)^2\, x_{t-2} + \ldots$$
$$= \alpha \sum (1-\alpha)^k (x_{t-k}) + (1-\alpha)S_0(x)$$

Thus, if the smoothing constant equals .2, then the current data point has weight .2. Previous observations have weights .16, .128, .1024, etc..

It is seen from the above equations that exponential smoothing always requires a prior estimate, S_{t-1}, to perform the update and find S_t. Brown [Ref. 6: p. 102] suggests using the simple average of the most recent N observations, or M_{t-1}, for the initial value S_{t-1}. In this study, the prior estimates are found in a similar manner. The functions developed by Tucker and Robinson [Refs. 1,2] and spelled out in detail in Robinson's thesis [Ref. 2: pp. 83-110] are used to find the estimates over N years of data. An optimal base estimate length N will be found for each aggregate using the data available. However, so that the results of this study may be compared with those of Robinson [Ref. 2], 3-year base estimates, corresponding to years 1977-1979, will be used. The empirical rates for 1980 are smoothed onto the base estimates to produce the updated estimates which can be validated on years 1981-1983 just as Robinson's estimates were [Ref. 2].

The exponential smoothing model best meets the criteria outlined in the previous section. It requires very little data to be carried from year to year as compared with the other candidates examined; only the last estimates obtained by the model need be saved. It is also very flexible to changes in the pattern of the data. When the smoothing constant is small, the function behaves like the average of a great deal of past data, whereas large values of the smoothing constant allow S(x) to respond quickly to changes in the attrition rate process [Ref. 6: p. 102]. Its accuracy should be greater than that of the moving average,

where the data is equally weighted, since exponential smoothing discounts past data. This will allow for more recent data to exert greater influence on the estimate, which is desirable because the near past generally represents the near future better than the distant past.

Therefore, since it fulfills the criteria better than all other candidates examined, exponential smoothing is the yearly updating model used in this work.

III. EXPONENTIAL SMOOTHING

A. GENERAL

Having chosen exponential smoothing for the model to update the attrition rates annually, the next step is to implement it. The exponential smoothing itself will be performed in transformed scale because the variances of the estimates are more stable in this scale. Therefore, both the base estimates $S_{i-1}(x)$ and observations x_i, which are needed to produce the updated estimate $S_i(x)$, will be calculated in the transformed scale.

When the loss data is collected by the Navy Personnel Research and Development Center, the cell (OF, LOS, grade) into which a loss is assigned is the cell to which the leaving officer belongs at his time of departure. For example, a first lieutenant with 3 years of service at the beginning of year i who completes his 4th year of service, is promoted to captain, and subsequently leaves the service before the beginning of year i+1 is classified as a loss from LOS 4 years in his OF and the grade of CAPT in year i. This type of loss classification demands the use of central inventory data [Ref. 7: p. 25]. For data grouped by years, like the Marine Corps manpower data we have, central inventory for year i is found by averaging beginning-of-year i stocks with beginning-of-year i+1 stocks for each cell. If beginning-of-year i+1 stocks are not available, because we are at the end of the data set or for whatever reason, the central inventory for year i is set equal to the beginning-of-year stocks for year i. This occurs in our data for year 1983, since 1984 is not available for our use. Additionally, if losses in year i are greater than the central inventory of year i, the central inventory is set equal to the losses to avoid the apparent inconsistency of losing more men than

you have on hand. Thus, in the above example, the officer would have been counted as a 1LT in his OF with 3 years of service in the inventory for the beginning of year i. The attrition rate computed then, is the central attrition rate, which is the number of losses in year i divided by central inventory for year i [Ref. 7: p. 25]:

$$m_i = y_i/N_i \tag{3.1}$$

The raw data compiled by the Navy Personnel Research and Development Center consists of losses and beginning-of-year inventories for the 12400 cells mentioned in the Introduction chapter. Both Major Tucker and Major Robinson [Refs. 1,2] aggregated this data into the four categories aviation (OF 75), combat support (OF 13, 25, 35), ground combat (OF 03, 08, 18), and other. The central inventory matrices were computed and the loss matrices compiled for the aviation, combat support, and ground combat aggregates for all years available (1977-1983 for Robinson and this work). The 'other' category was not examined by Tucker nor Robinson, nor will it be here. The best summary of the data manipulation programs producing the central inventory and loss matrices is found in Appendix D of Major Robinson's thesis [Ref. 2: pp. 83-104]. In order to be consistent with the analyses done in those two works, only the grades of first lieutenant (1LT) and lieutenant colonel (LTCOL) will be examined in this pilot study. Like Tucker and Robinson [Refs. 1,2], all 31 possible lengths of service years will be included. Additionally, all six estimators explored by Robinson [Ref. 2] and discussed in the Introduction chapter of this study will undergo exponential smoothing and have their forecasting abilities compared. Thus the total number of loss rate estimates that will ultimately be computed is 25284, corresponding to 6 estimators, 7 operational fields,

31 lengths of service, and 2 grades. The analysis will be broken down by grade and aggregate, as in Tucker and Robinson's works [Refs. 1,2], thereby setting up 6 study blocks:

1) Aviation 1LTS
2) Aviation LTCOLS
3) Combat Support 1LTS
4) Combat Support LTCOLS
5) Ground Combat 1LTS
6) Ground Combat LTCOLS .

The base estimate S_{i-1} and empirical estimate x_i are found for all six estimates, 31 lengths of service, and occupational fields within a study block by using APL function ESTIM (see Figure A.3), developed by Major Robinson [Ref. 2: p. 105]. The updated estimate, still in transformed scale, is found by applying exponential smoothing equation 2.1 presented in the preceding chapter. The predicted rates are then validated against actual transformed figures from the years that follow. How this is done is the subject of the next section.

B. FIGURES OF MERIT

The figures of merit (FOM) used by Robinson in his thesis [Ref. 2] will also be used here. The basis for the decisions concerning the finding of the optimal smoothing constants and base period lengths will be the figures of merit in transformed scale because the data is so much more well-behaved therein. The original scale figures of merit will also be calculated and reported in Chapter IV. The values of α used to calculate the original scale figures of merit will be the optimal values of α found in transformed scale.

We would hope, therefore, that the value of the smoothing constant, α, corresponding to the smallest figure of merit in transformed scale is very close to that

corresponding to the smallest FOM in original scale. Fortunately, this is almost always the case. Two exceptions are the transformed aggregate estimators for aviation 1LTS and ground combat 1LTS, which turn out not to be very important due to the overall lackluster performance the transformed aggregate estimator turns in throughout the analyses. More notably, however, the MLE for combat support 1LTS, ground combat 1LTS, and ground combat LTCOLS have serious discrepancies between values of α producing minimum figures of merit in transformed and original scales. These differences cast serious doubts upon the correctness of using the transformed scale optimal α as the α value in the original scale for maximum likelihood estimation. See Appendix C for the details of this analysis.

In transformed scale, the data is approximately distributed normally with a stable variance (see Appendix B). Thus, a good means of comparison between estimators is a sum of squares error (SSE) computation. The sum of squares error is defined as:

$$\text{SSE} = \sum (\text{actual-predicted})^2 \qquad (3.2)$$

where "actual" and "predicted" are the actual and predicted values for the transformed attrition rate figures for one of the validation years (VY). APL function RISKT calculates the figures of merit for transformed scale using the above equation (see Figure A.9).

The original scale does not have a normal distribution nor a stable variance for the estimates. The SSE would therefore be inappropriate for the FOM calculations in original scale. Robinson [Ref. 2: p. 28] sugggested the use of a chi-square statistic for use to compare estimators in original scale. The chi-square statistic is:

$$\mathrm{FOM}(t) = \sum \; (y_{it}-N_{it}p_i)^2/N_{it}p_i(1-p_i), \text{ for all } i \qquad (3.3)$$

where y_{it} and N_{it} are the losses and central inventory counts for the i^{th} cell in the t^{th} validation year, and p_i is the inverse transform of the estimate for the i^{th} cell calculated in transformed scale (see Appendix B). APL function RISKO (see Figure A.8), which computes the original scale figures of merit, screens out cells with p_i values of 0 and 1 to prevent the above denominator from having a value of zero. Both RISKT and RISKO find the figures of merit for all of the validation years available.

With both of these figures of merit, the smaller the FOM, the better the estimates produced. Figures of merit are summed over all of the operational fields and lengths of service within a study block, thereby producing a single number for necessary comparisons.

C. FINDING THE SMOOTHING CONSTANTS

As mentioned in the preceding paragraph, it is the smallest figure of merit which we seek. Therefore, the *smoothing constant*, α, which produces the smallest FOM for an estimator is the optimum α for that estimator.

In order for the weighting scheme implied by exponential smoothing to make sense, α must be between 0 and 1 inclusive. An estimate with a value of α equal to 0 places no weight on the current empirical data point and all of the weight on the previous estimate; thus, this value is simply the estimate produced by Robinson [Ref. 2]. Conversely, an estimate produced with a value of $\alpha = 1$ is simply the empirical estimate from the most recent year's data.

An APL function was developed to produce the figures of merit in both transformed and original scale for all validation years available and all six estimators for a study block for all values of alpha between 0 and 1 by a specified step size, which in this study is .02. This function,

ALPHAHAT, calculates the base estimates using function ESTIM, the empirical data points using function XFOUR (for a three-year base; XFIVE and XSIX for four- and five-year bases), which calls ESTIM, and smooths them into the new estimates using equation 2.1. See Appendix A. Arrays of figures of merit are produced, which can then be analyzed and the optimum α found for each estimator and study block.

D. LENGTH OF BASE PERIOD

The length of the base period is important in producing the loss rate estimates. If the environment is stable over a long period of time, a long base period is preferred to smooth out random fluctuations in the data. If the environment is in a constant state of turmoil, the base period length should be rather small, and frequent updates of the base are needed.

To provide a basis for comparison with the results obtained by Robinson [Ref. 2], a three-year base will be used in this study, with the fourth year empirical data smoothed onto it to produce the attrition rate estimates. Thus, years 1977-1980 are used for computation of the estimates, and years 1981-1983 are used for validation, as in Robinson's work [Ref. 2].

It should be noted that the choice of three years as the length of the base period may not be "optimal," that is, it may not provide estimates as accurate as those of a four- or five-year base period. Therefore, a short comparison of base estimate lengths three, four, and five years is presented below. The MOE used is transformed FOM for forecasted rates for validation year one. Although we are more concerned with the estimators' ability to estimate rates two years in advance, this MOE is chosen so that a 5-year base, which can only forecast for 1983, may be included.

To find the figures of merit, the optimal α, which is the value of α producing the smallest FOM for validation

year one, is found for each base year length and study block. The resulting figures of merit are then compared to determine the "best" base period length for an aggregate.

It should be noted that with the limited number of data years we have to work with, the conclusions drawn by this short analysis cannot be generalized to future years. The length of base period problem should be examined in more detail as more data becomes available.

TABLE 1

DETERMINATION OF OPTIMAL BASE PERIOD -- COMBAT SUPPORT

FIRST LIEUTENANTS

Estimator	3-Year Base		4-Year Base		5-Year Base	
	α	FOM81	α	FOM82	α	FOM83
AGG ORIG	.50	1.863	1	1.669	.62	1.494
AGG TRANS	0	4.086	0	4.737	0	4.264
MLE	.54	2.402	.58	2.409	.48	1.894
TSCA	.52	1.610	.62	1.877	.50	1.136
JS	.78	1.678	.80	1.817	.58	1.153
LTJS	.58	1.641	.66	1.861	.52	1.133

LIEUTENANT COLONELS

Estimator	3-Year Base		4-Year Base		5-Year Base	
	α	FOM81	α	FOM82	α	FOM83
AGG ORIG	.66	1.478	1	1.441	1	1.362
AGG TRANS	1	3.372	1	2.789	0	2.603
MLE	.36	1.859	.50	1.625	.40	1.638
TSCA	.38	1.024	.72	1.209	.60	.980
JS	.36	.935	.76	1.124	.64	.934
LTJS	.36	.968	.74	1.143	.62	.956

Note: Figures of Merit are those for transformed scale.

1. Combat Support

Interestingly, there is an increase in the minimum FOM from a three- to a four-year base, with the five-year base minimum FOM being smaller than both of them for the TSCA, JS, and LTJS estimators for both grades, as well as AGG TRANS and MLE 1LTS. The other 4 estimator-grade combinations show a strictly decreasing trend across all 3 base

period lengths. Thus, from Table 1, it appears that a 5-year base would be best to use for the combat support aggregate, since all transformed figures of merit are smallest for this base period length.

TABLE 2

DETERMINATION OF OPTIMAL BASE PERIOD -- GROUND COMBAT

FIRST LIEUTENANTS

Estimator	3-Year Base		4-Year Base		5-Year Base	
	α	FOM81	α	FOM82	α	FOM83
AGG ORIG	.58	2.976	1	3.066	1	3.933
AGG TRANS	0	19.842	0	20.923	0	24.811
MLE	.72	1.647	.82	2.127	.80	2.909
TSCA	.8	1.324	1	1.668	.90	2.559
JS	.92	1.460	1	1.664	.96	2.631
LTJS	.86	1.348	1	1.609	.94	2.573

LIEUTENANT COLONELS

Estimator	3-Year Base		4-Year Base		5-Year Base	
	α	FOM81	α	FOM82	α	FOM83
AGG ORIG	1	3.540	1	2.698	1	2.413
AGG TRANS	0	14.393	0	12.481	0	12.653
MLE	.42	1.389	.20	1.982	.40	2.250
TSCA	.58	1.510	.58	1.904	.70	1.922
JS	.58	1.505	.60	1.832	.70	1.884
LTJS	.56	1.494	.58	1.881	.70	1.899

Note: Figures of Merit are those for transformed scale.

2. <u>Ground Combat</u>

From Table 2, an increase is seen from both 3- to 4-year bases and from 4- to 5-year bases in all cases except AGG ORIG and AGG TRANS for LTCOLS. Overall, it appears that a 3-year base period produces the smallest figures of merit for validation year one, and is therefore "optimal" for this aggregate and data set.

The large values of α for first lieutenants are noteworthy. They may represent the inherent variability of the process, or they may indicate that the attrition process for this aggregate is in an almost continuous state of

change from year to year. The reasons are unclear at this time, and further analysis with more data is needed to determine whether or not this trend persists.

TABLE 3

DETERMINATION OF OPTIMAL BASE PERIOD -- AVIATION

FIRST LIEUTENANTS

Estimator	3-Year Base α	3-Year Base FOM81	4-Year Base α	4-Year Base FOM82	5-Year Base α	5-Year Base FOM83
AGG ORIG	1	3.976	1	4.253	.44	4.736
AGG TRANS	0	206.877	0	218.857	0	216.057
MLE	.8	2.177	1	4.287	.36	1.502
TSCA	.98	1.243	1	4.512	.38	1.176
JS	1	1.402	1	4.494	.38	1.277
LTJS	1	1.386	1	4.518	.38	1.261

LIEUTENANT COLONELS

Estimator	3-Year Base α	3-Year Base FOM81	4-Year Base α	4-Year Base FOM82	5-Year Base α	5-Year Base FOM83
AGG ORIG	1	7.898	1	10.959	1	6.992
AGG TRANS	1	24.733	0	31.473	0	27.999
MLE	.88	3.379	.66	6.637	.50	3.208
TSCA	1	3.341	.94	6.325	.7	3.698
JS	1	3.302	.96	6.311	.72	3.719
LTJS	1	3.321	.94	6.377	.7	3.684

Note: Figures of Merit are those for transformed scale.

3. Aviation

From Table 3, it is seen that for aviators, the 4-year base period has a higher FOM than the other two candidates in all cases except for AGG ORIG 1LTS, for which the FOM for a 4-year base length is in the middle of the three values. Thus, the choice for the "optimal" base period length centers on the 3- and 5-year candidates. For 1LTS, the minimum FOM for the 5-year base is smallest for MLE, TSCA, JS, and LTJS, while the 3-year base is "best" for both of the aggregate estimators. For LTCOLS, the 5-year base has the smallest minimum FOM only for AGG ORIG and MLE, while the 3-year base excels for the other four. With six

estimators being "best" for both the 3- and 5-year base period lengths, an aggregate-wide analysis results in a stalemate. Thus, the analysis is broken down by grade, with the conclusions being that the "optimal" base period lengths are 5 years for aviation 1LTS, and 3 years for LTCOLS.

The α values in this table, shifting from near 1 for a 3- and 4-year base to around .38 for 1LTS and .7 for LTCOLS for the 5-year base should be noted; they indicate a radical change in the data. This change comes about, in part, because of the initiation of Aviation Officer Continuation Pay (AOCP) in 1981. As explained by Major Tucker [Ref. 1: p. 18], AOCP provides a bonus per year to aviation officers which in turn obligates continued service. The program was applied to all ranks provided the individual met certain active duty flight status requirements. This action had its desired effect on retaining aviation officers, according to the analysis of Major Tucker [Ref. 1: p. 18].

In this analysis, the effect of AOCP is that data from the pre-1981 era is not relevant to post-1981 data, thereby producing α values of 1.0 for the 3- and 4-year base period lengths for aviation. This means that the estimates producing the smallest figures of merit are simply the prior year's empirical estimates, which makes sense when the base years include all pre-1981 data, as they do for the 3- and 4-year bases. The values of α in the .7 range for 3 of the 6 estimators for lieutenant colonels and the 5-year base indicate that theoretically, 1982 data should be given a weight of .7, 1981 data a weight of (.7)(.3) = .21, and 1977-1980 data a collective weight of $(.7)(.3)^2$ = .09 (see Equation 2.2). This is also consistent with the AOCP program; the exponential smoothing model is beginning to build a base of its own beginning in 1981. This is also seen to be happening to the 4 estimators introduced by Robinson

[Ref. 2] for first lieutenants; the model is constructing its own base. The exponential smoothing model really shows its flexibility and overall value in its predictions for the aviation aggregate.

E. STABILITY OF THE SMOOTHING CONSTANTS

The stability of the smoothing constants is important to the validity of the exponential smoothing model. This is because if the same value of α produces optimal results in the short run, then the exponential smoothing model is a good one to use. We say "in the short run" because changes in conditions over time that affect the data will demand updates in α as well as the base period used.

Despite the relatively short base periods forced upon us by a lack of data, the smoothing constants found for each study block appear to be rather stable, i.e., there are no wild fluctuations in the α which produces the minimum transformed FOM. The few exceptions will be identified in the analysis that follows. The stability of the smoothing constants will be demonstrated in two ways: within base period lengths, and between base period lengths.

1. Within Base Period Lengths

The stability of α within a base period length is measured by how much α varies in producing minimum figures of merit for all of the validation years (VY) available. The analysis of the stability of α within base period lengths, therefore, is restricted to the 3- and 4-year bases, since the 5-year base can only be validated on one year, 1983, in our data set. In addition, the two aggregate estimators are excluded from this and the following (between base period lengths) graphical analyses because of the consistently poor performance they show as compared to the other four estimators as measured by the figures of merit in transformed space seen in Tables 1, 2, and 3. An analysis of the stability of α for these two estimators reveals that

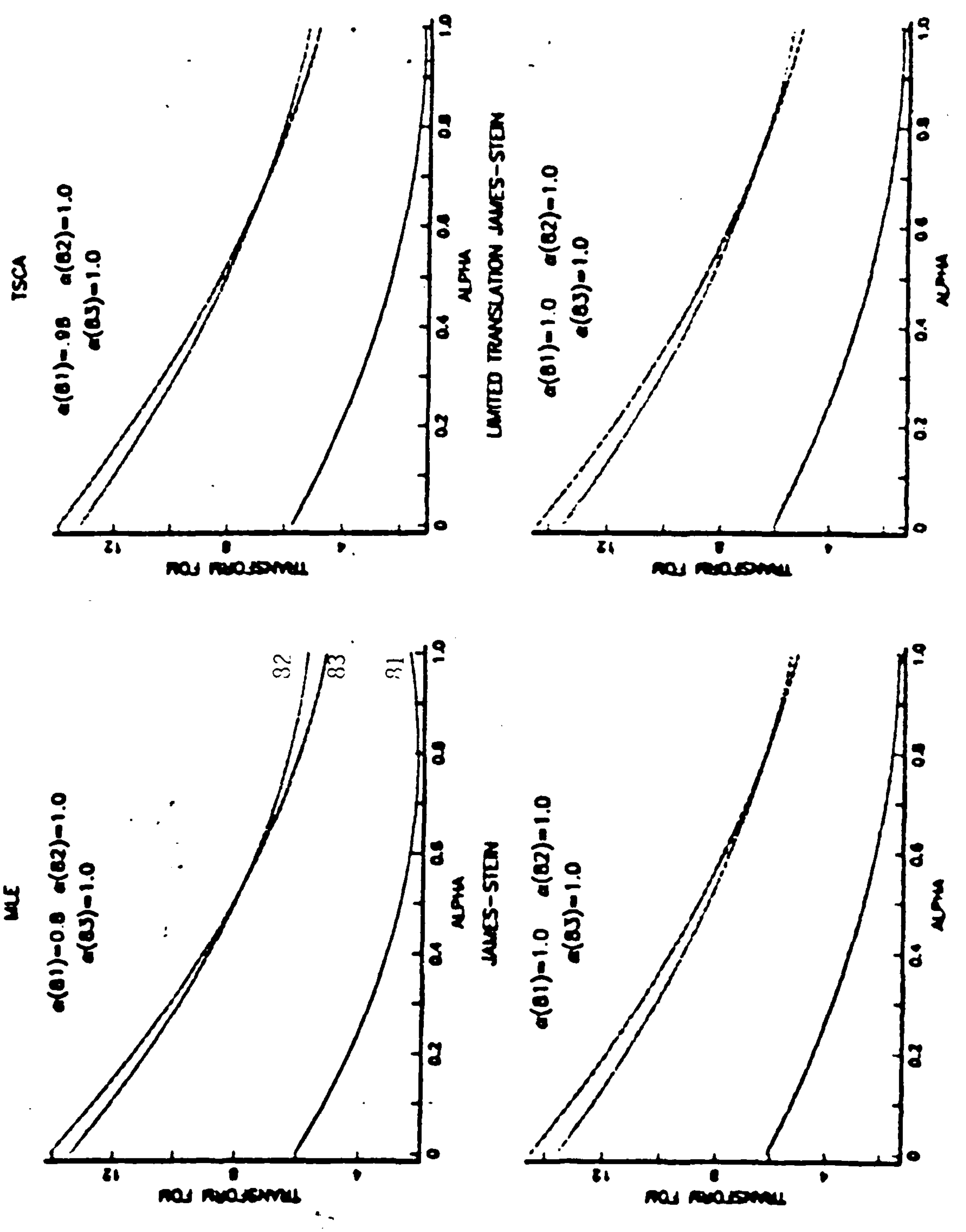

Figure 3.1 Within 3-Year Base Stability -- Aviation 1LTS

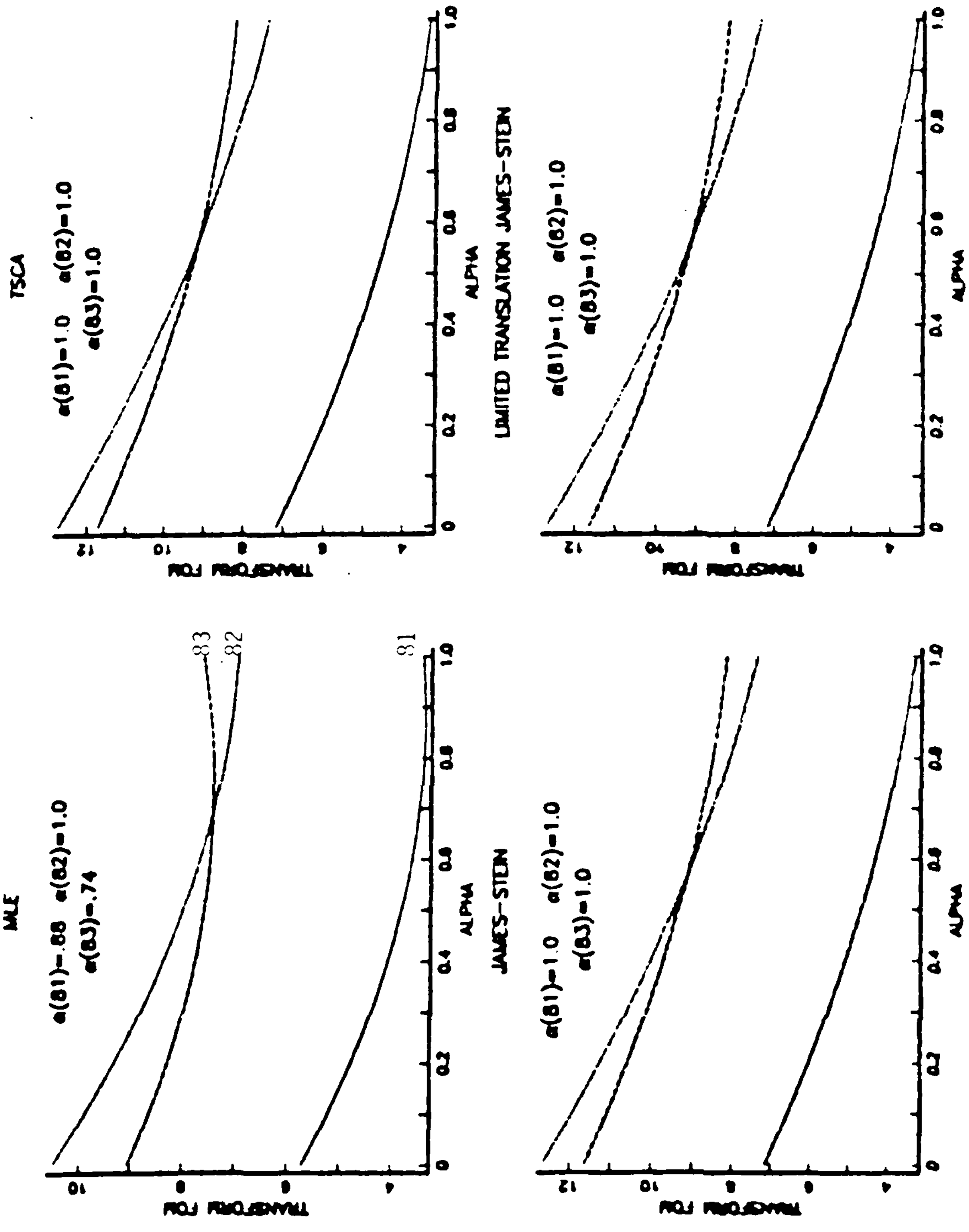

Figure 3.2 Within 3-Year Base Stability -- Aviation LTCOLS

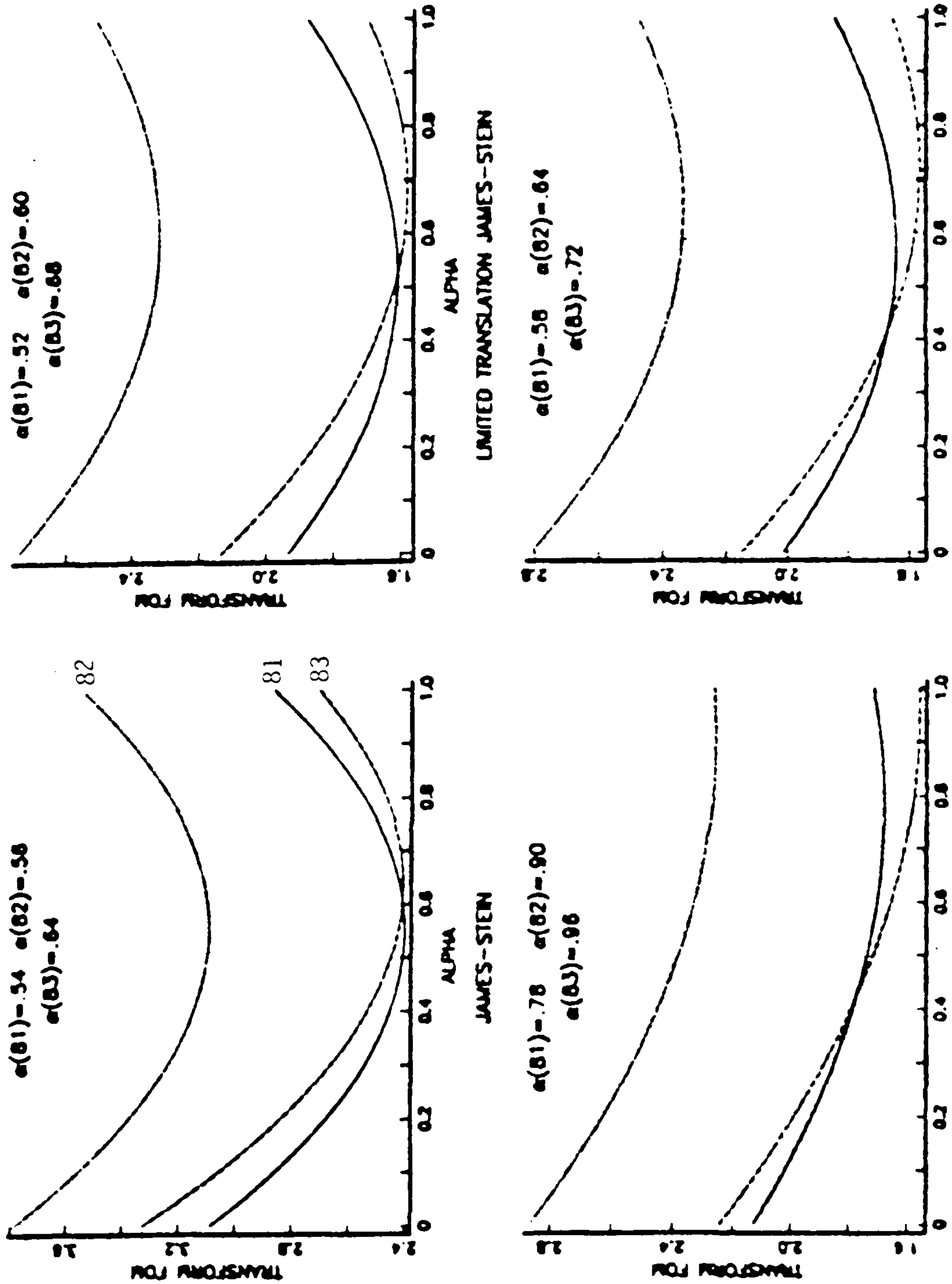

Figure 3.3 Within 3-Year Base Stability -- Cbt. Spt. 1LTS

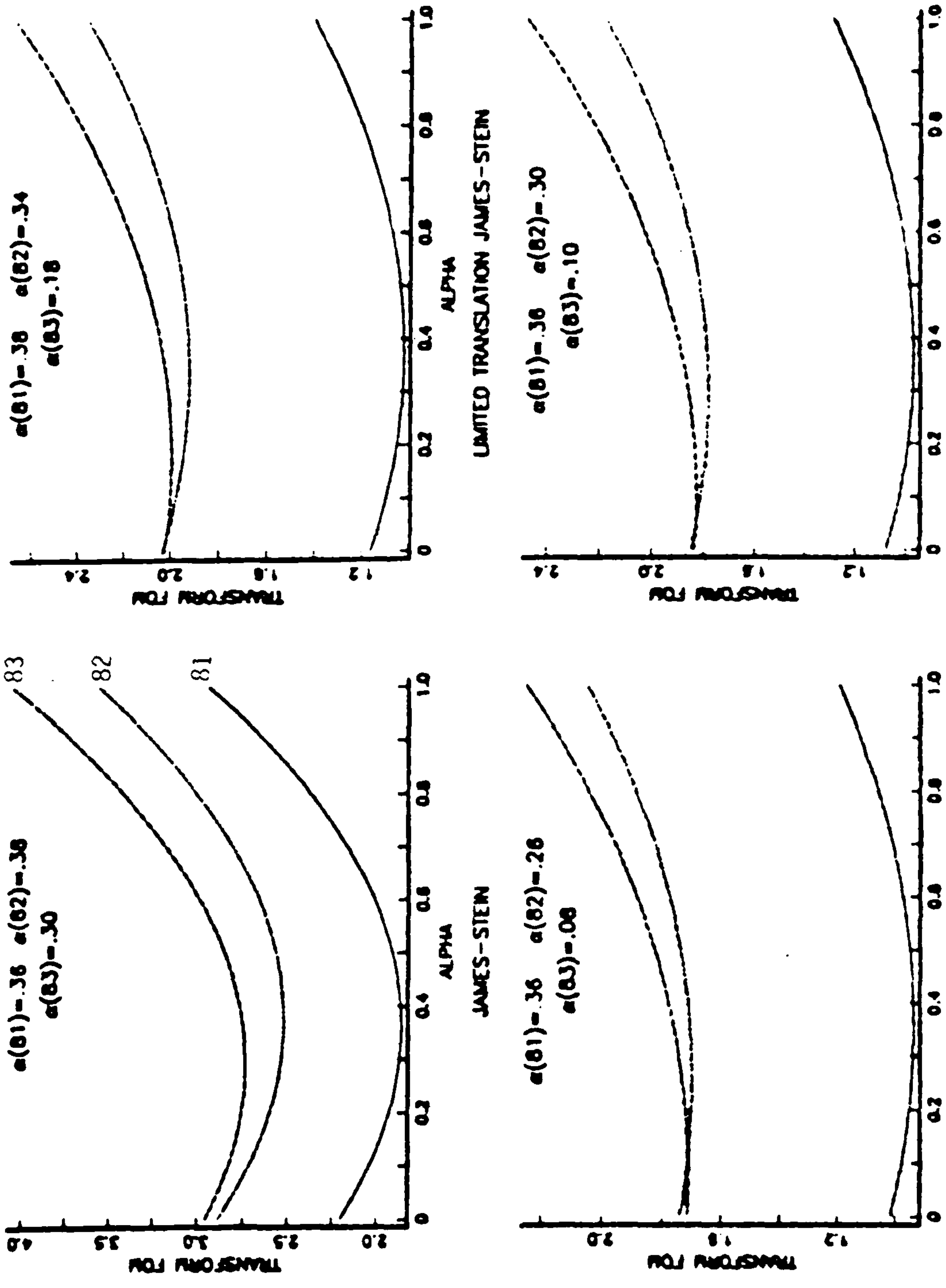

Figure 3.4   Within 3-Year Base Stability -- Cbt. Spt. LTCOLS

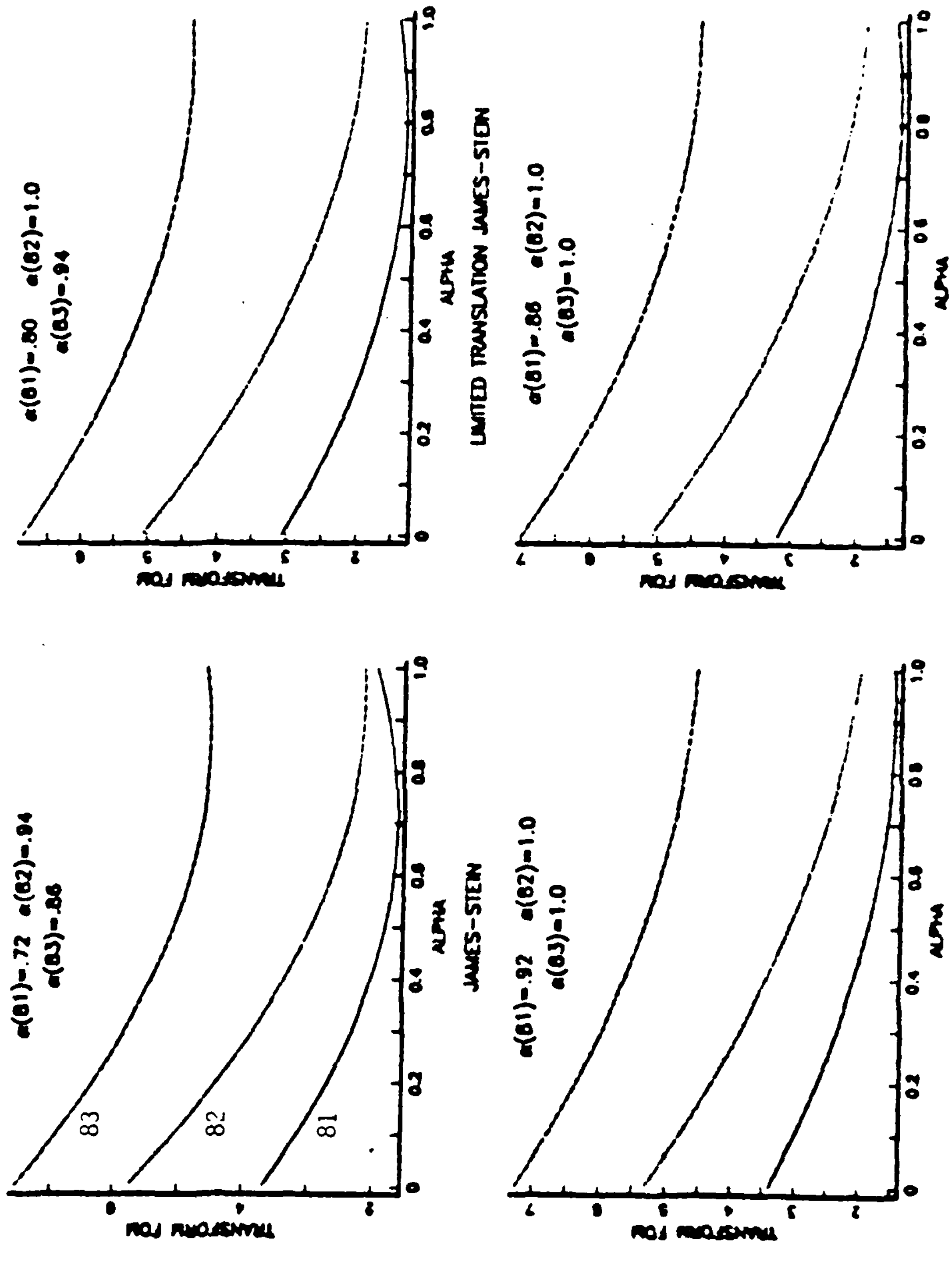

Figure 3.5 Within 3-Year Base Stability -- Grd. Cbt. 1LTS

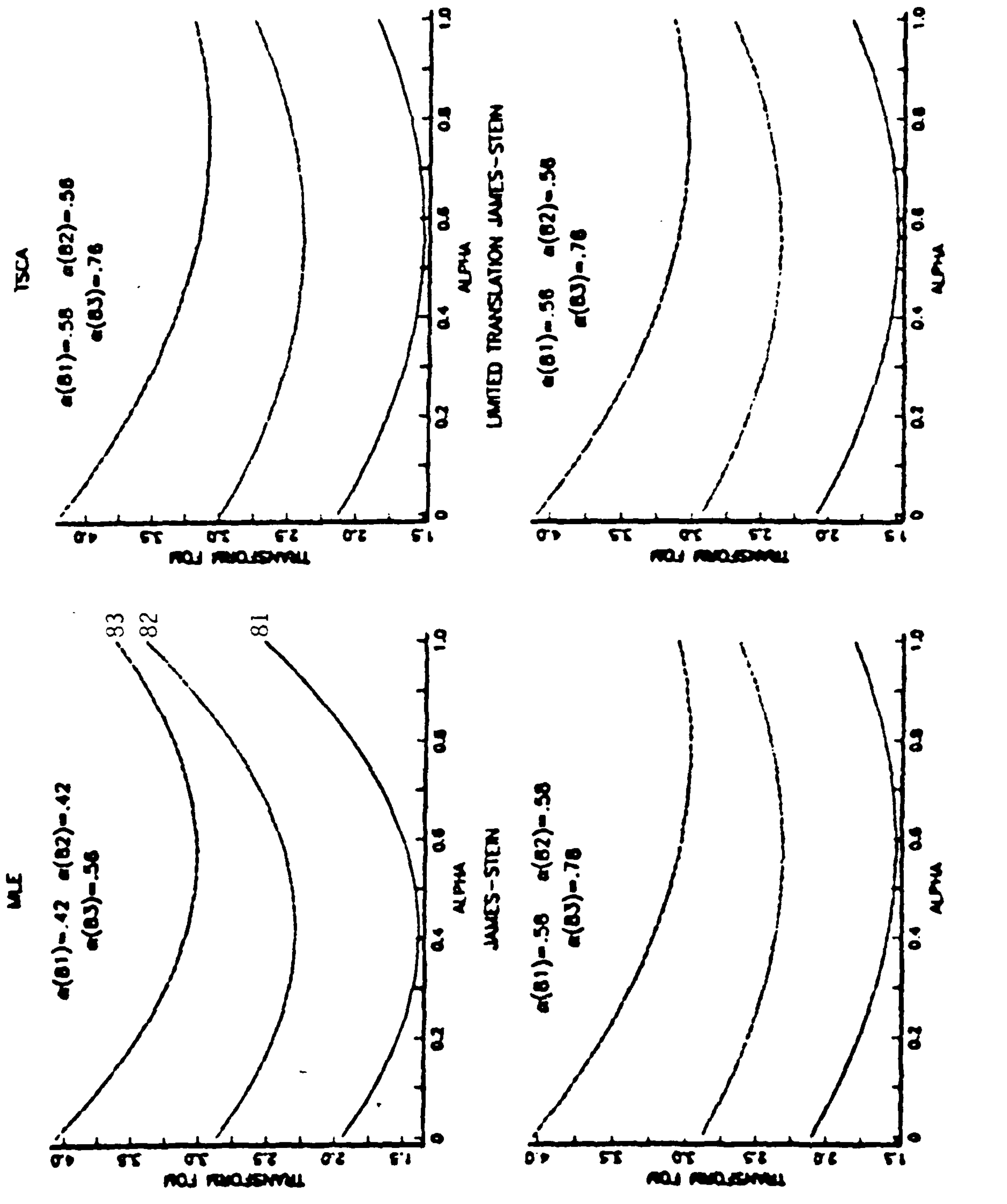

Figure 3.6 Within 3-Year Base Stability -- Grd. Cbt. LTCOLS

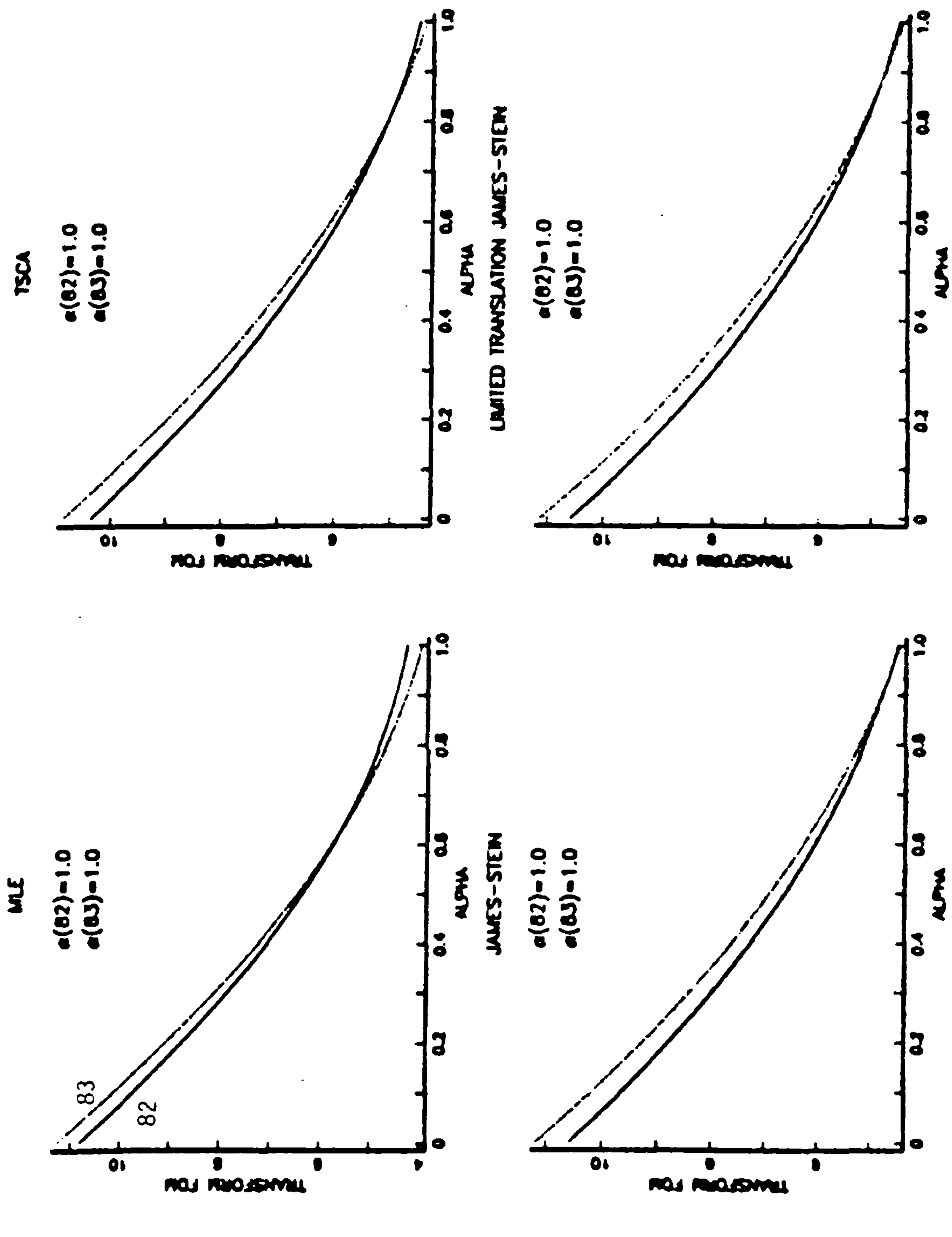

Figure 3.7   Within 4-Year Base Stability -- Aviation 1LTS

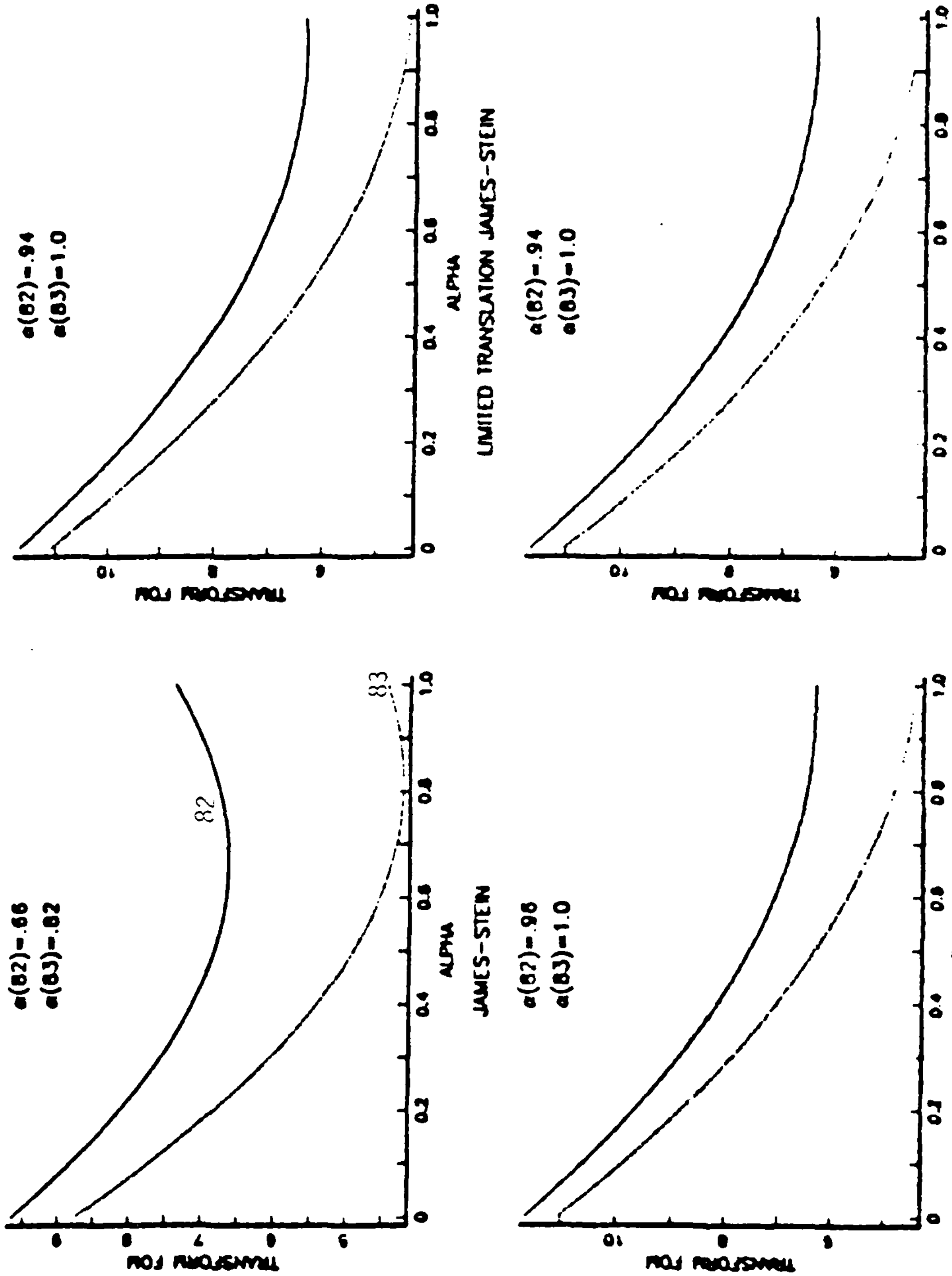

Figure 3.8 Within 4-Year Base Stability -- Aviation LTCOLS

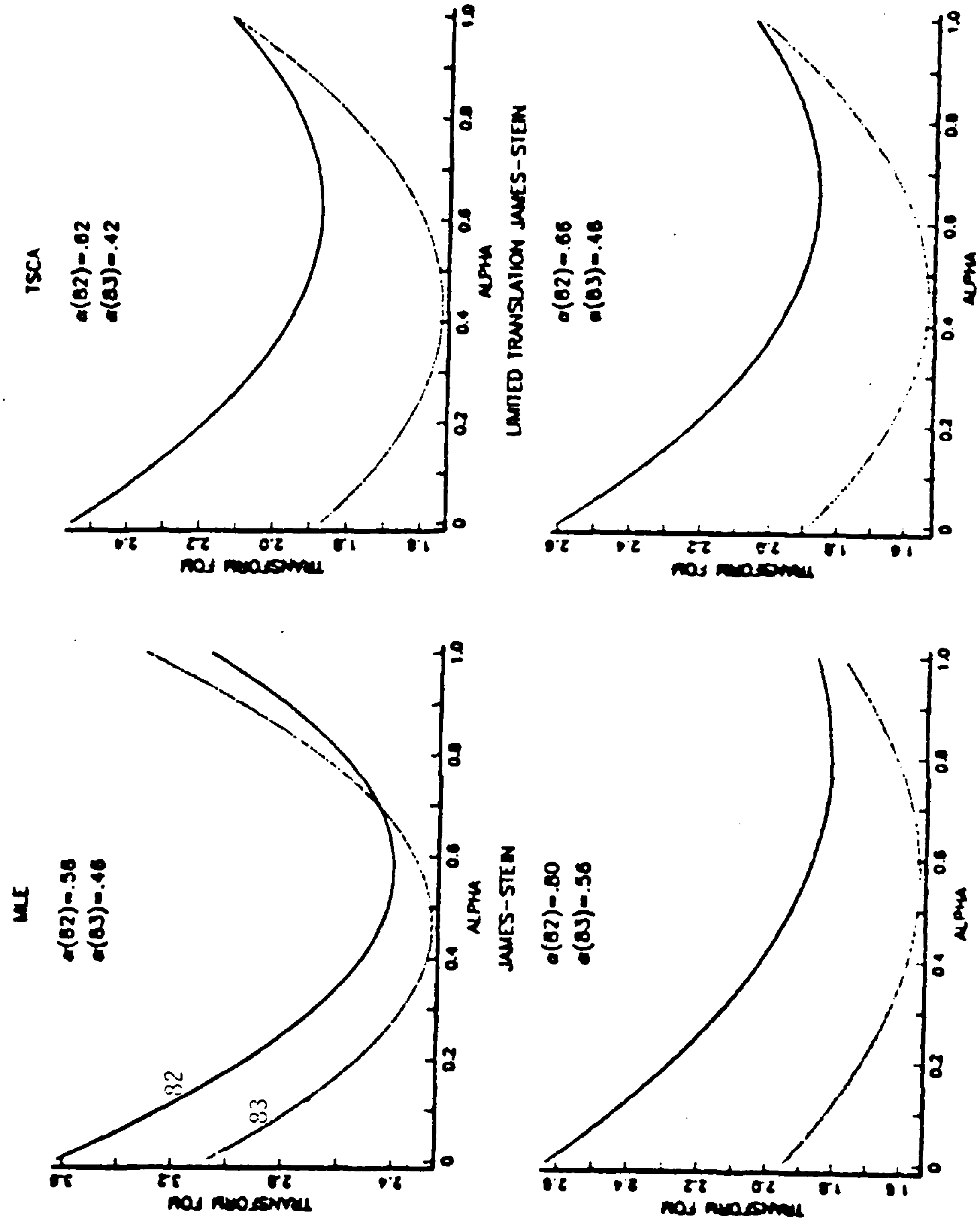

Figure 3.9 Within 4-Year Base Stability -- Cbt. Spt. 1LTS

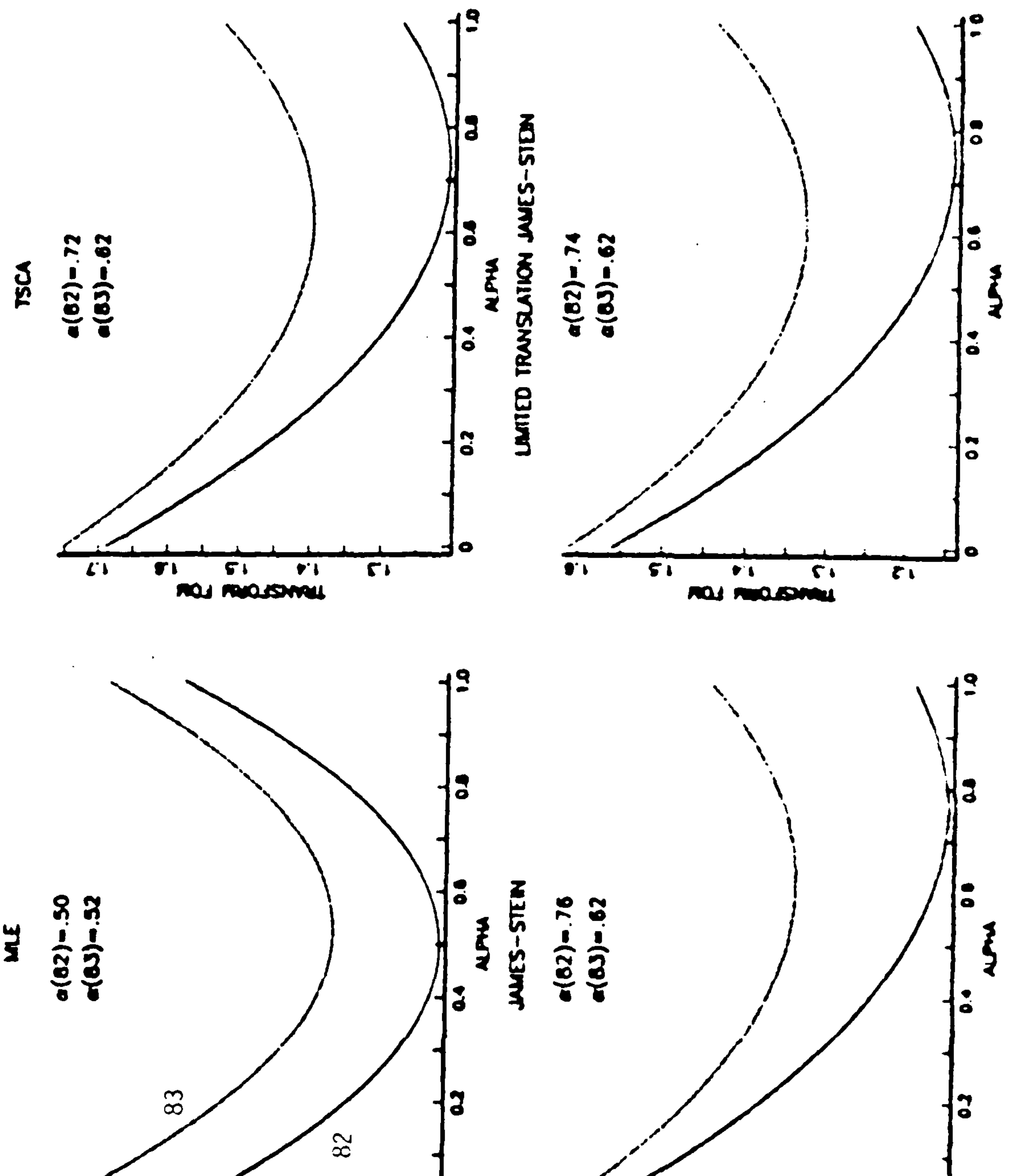

Figure 3.10 Within 4-Year Base Stability -- Cbt. Spt. LTCOLS

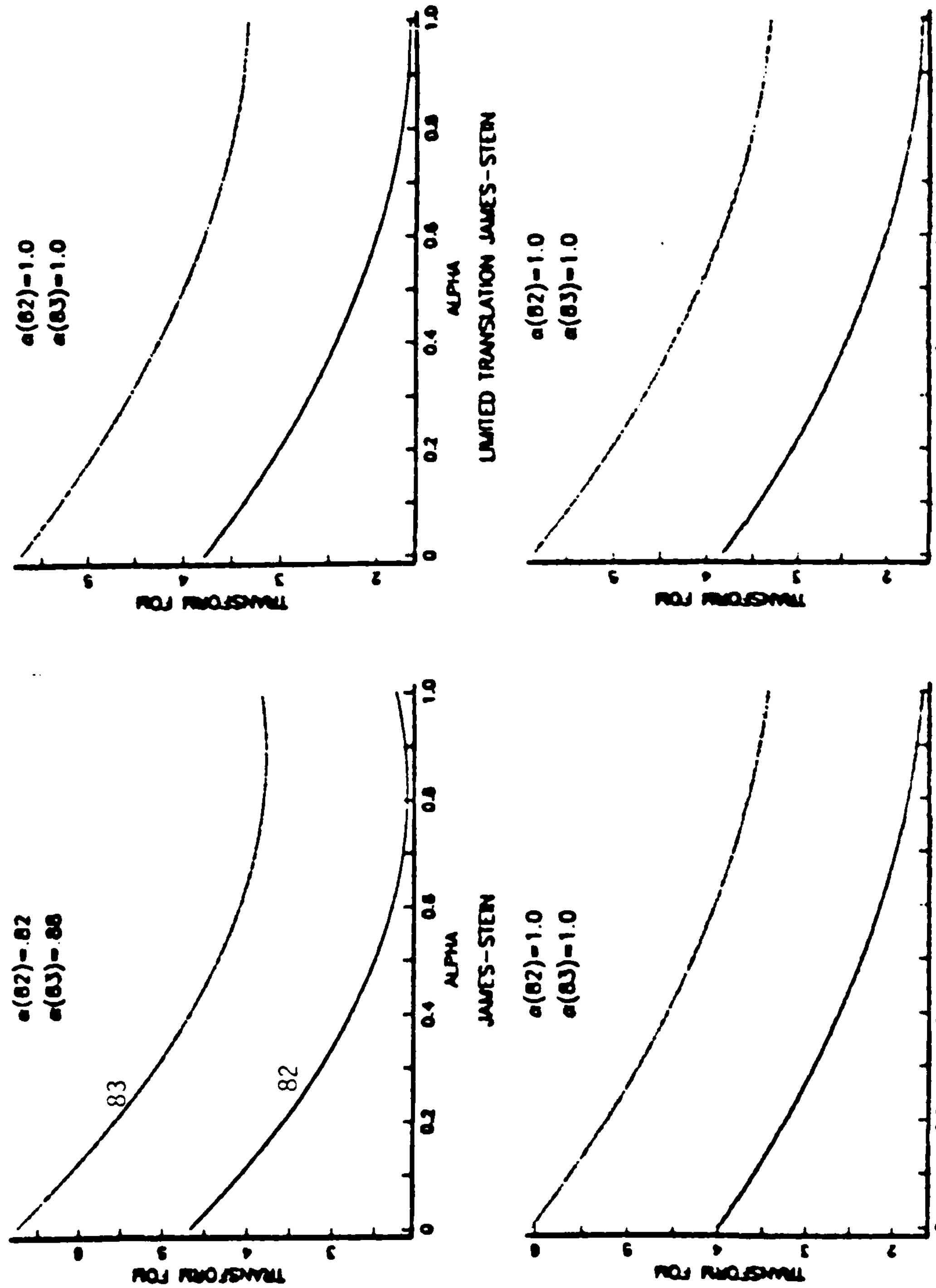

Figure 3.11 Within 4-Year Base Stability -- Grd. Cbt. 1LTS

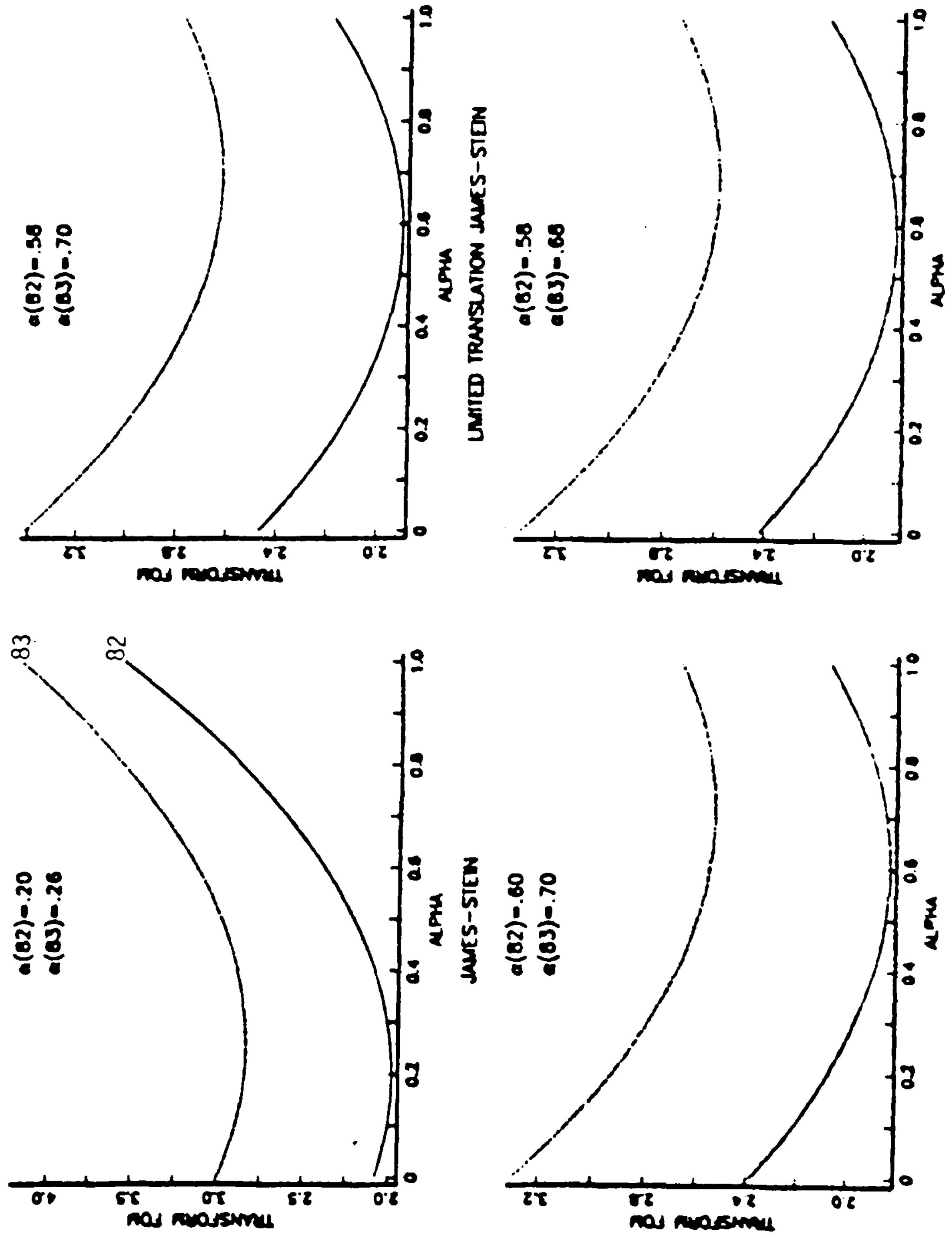

Figure 3.12   Within 4-Year Base Stability -- Grd. Cbt. LTCOLS

the AGG TRANS estimator shows perfect consistency in its optimal α values in all cases, and the AGG ORIG α values are also very stable in most cases with no notable exceptions.

Note that when reading the graphs for the 3-year base period length, the solid line is for validation year 1, 1981, the dot-dashed line is for validation year 2, 1982, and the dashed line is for validation year 3, 1983. The graphs for the 4-year base period length have a solid line for validation year 1, 1982, and a dot-dashed line for validation year 2, 1983. The year corresponding to each type of line is seen on one of the graphs, usually the "MLE" graph, on each page.

At first glance, it would appear that these twelve graphs show some study cells to have very stable α values while others have α values that vary greatly. Before any conclusions are drawn, however, attention must be focused on the scales of the Y-axes. In almost all cases, these scales are so small that they render the curves practically flat over the interval covered by the range of optimal α values. This indicates stability in α despite the seen difference in optimal α values since any α value in the range of optimal α values will produce a figure of merit very close to the optimal FOM.

A measure of the disparity caused by the differing α values is the maximum percent differential of figures of merit produced in the range of optimal α values over all validation years. This error is measured for each validation year, estimator, and study block by subtracting the minimum FOM produced in the range of α from the the maximum FOM produced and then dividing by the minimum FOM. The maximum of these over the validation years for an estimator and study block is the error measure.

For these figures, this value is generally very small, usually between 0 and 2 percent. For a 3-year base

period length, however, there are some estimator-study block combinations with errors greater than 5% (chosen arbitrarily to be an acceptable tolerance), and these are:

1) aviation 1LTS MLE with 15.84% error
2) combat support LTCOLS JS with 5.24% error
3) combat support LTCOLS LTJS with 5.78% error
4) ground combat 1LTS MLE with 12.87% error
5) ground combat 1LTS TSCA with 8.08% error
6) ground combat LTCOLS MLE with 5.75% error

This implies that only 6 of the 24 α values studied for the 3-year base period may vary too much within validation years for that base period length.

For a 4-year base-period length even better results are obtained. The maximum percent differential is the 4.2% posted by the James-Stein estimator for combat support 1LTS, which also had the largest gap in α values, .24. In fact, outside of that study group, only one estimator-study block combination, MLE for aviation LTCOLS, has a percent differential of more than 1.1%. Thus, Figures 3.1 to 3.12 apparently show that α is stable within base period lengths for both cases we can study using the data set.

2. <u>Between Base Period Lengths</u>

The stability of α between base period lengths is measured by how much α varies in producing minimum figures of merit for equivalent validation years. The analysis will therefore observe optimal α for validation year 1 using a 3-, 4-, and 5-year base period length, as well as validation year 2 using a 3- and 4-year base period length.

Note that when reading the graphs for validation year one, the solid line is for the 3-year base, the the dot-dashed line is for the 4-year base and the dashed line is for the 5-year base. The graphs for validation year two have a solid line for the 3-year base and a dashed line for the 4-year base. The year corresponding to each type of

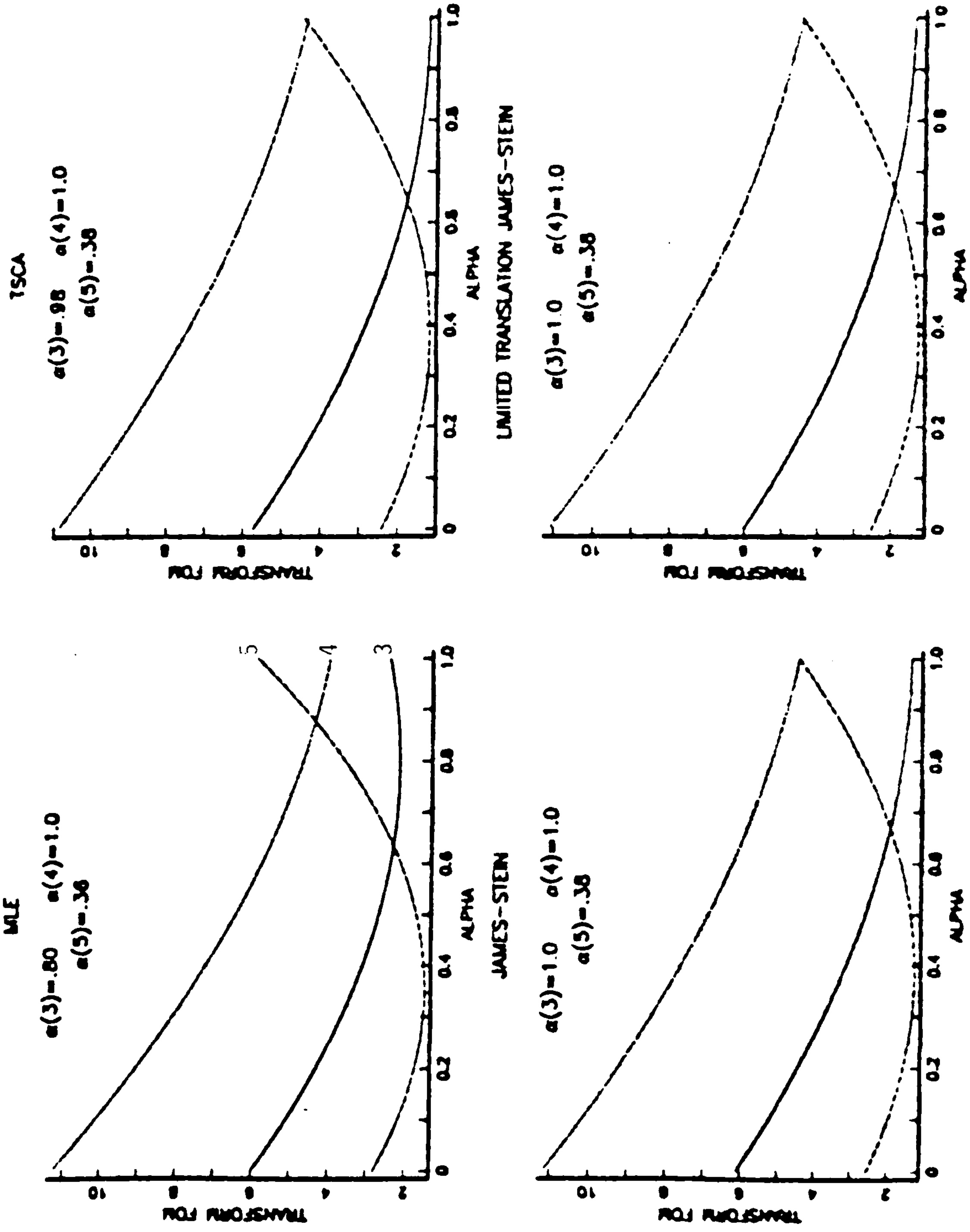

Figure 3.13 Between VY One Stability -- Aviation 1LTS

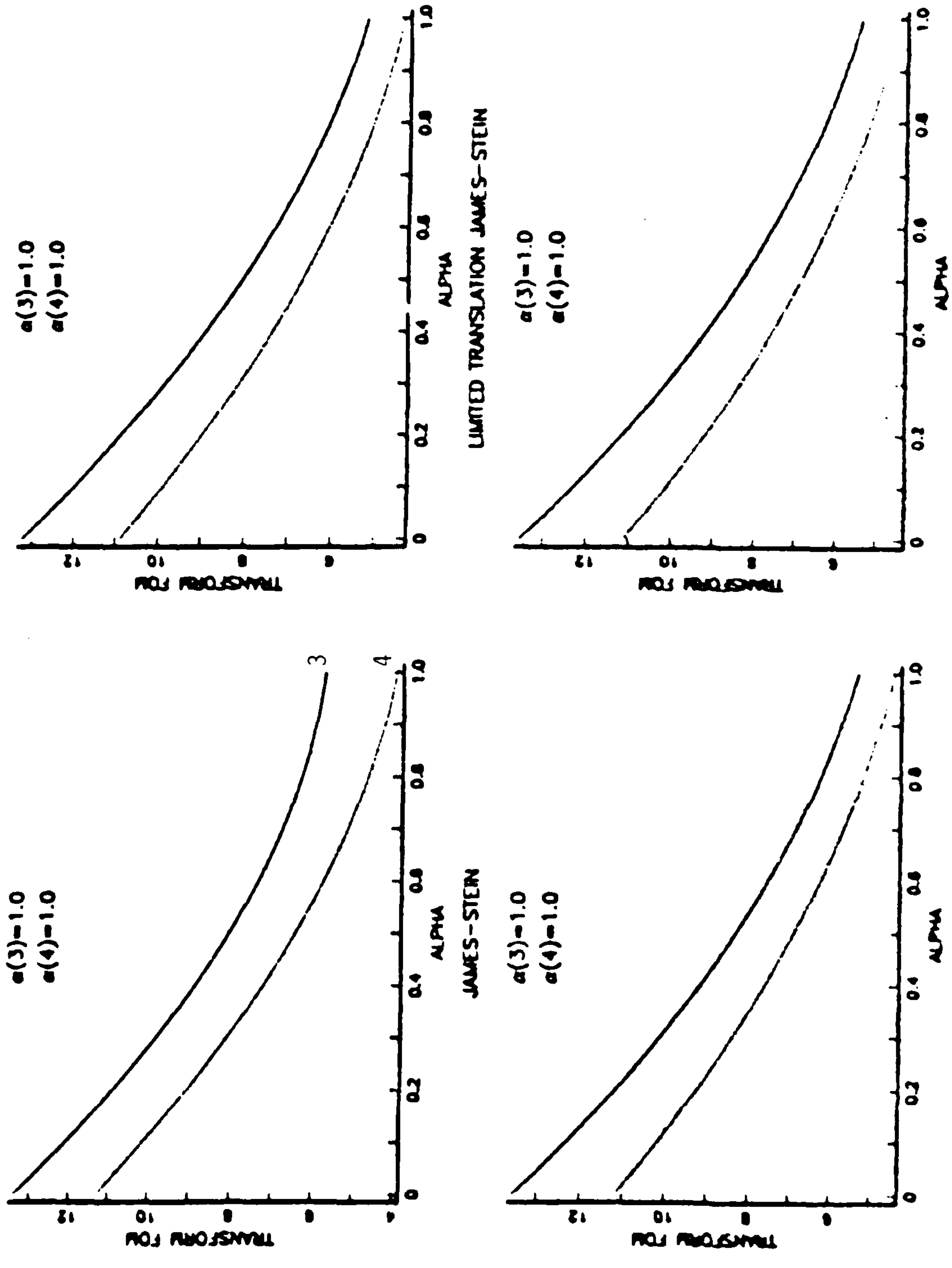

Figure 3.14 Between VY Two Stability -- Aviation 1LTS

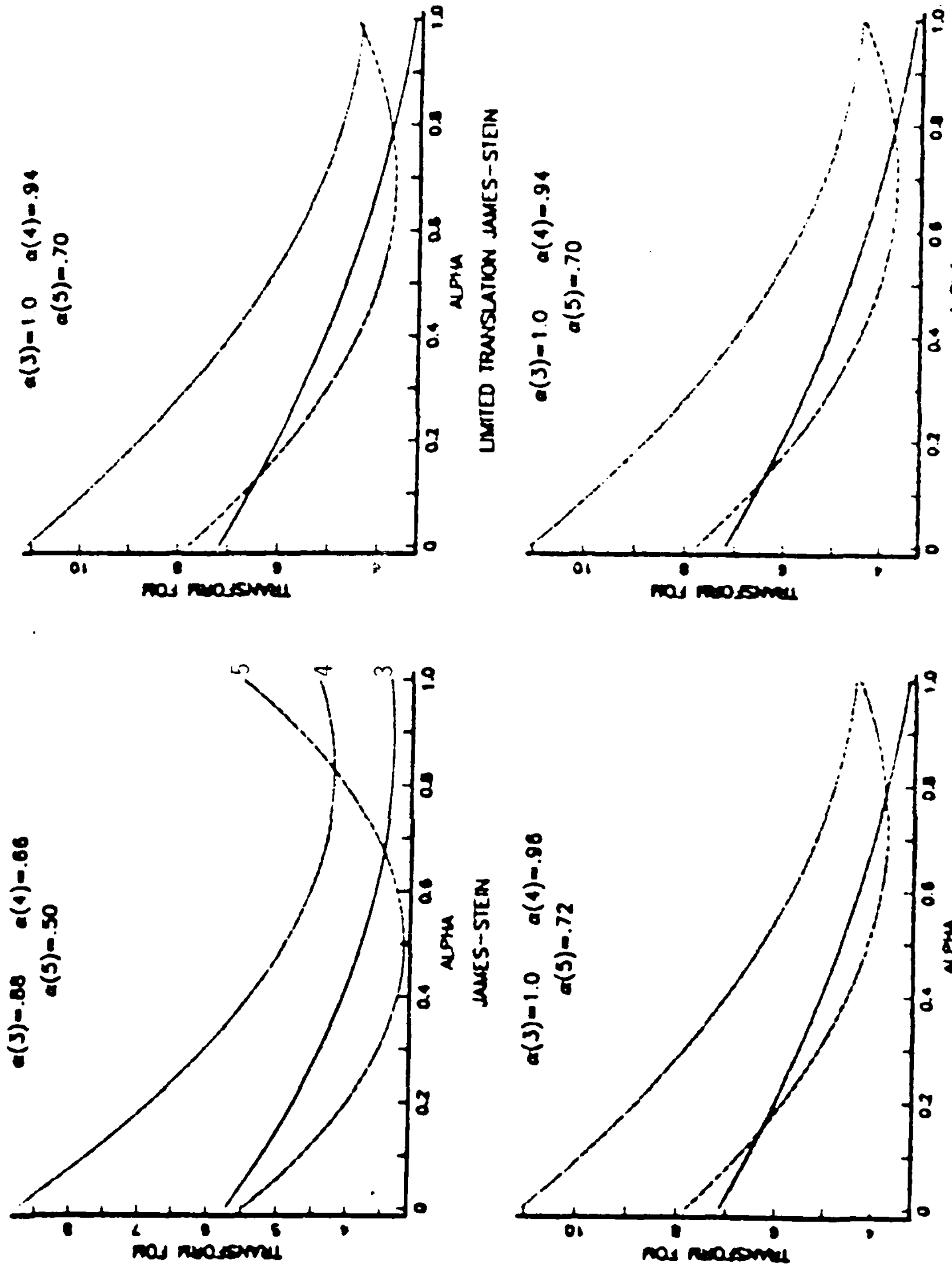

Figure 3.15 Between VY One Stability -- Aviation LTCOLS

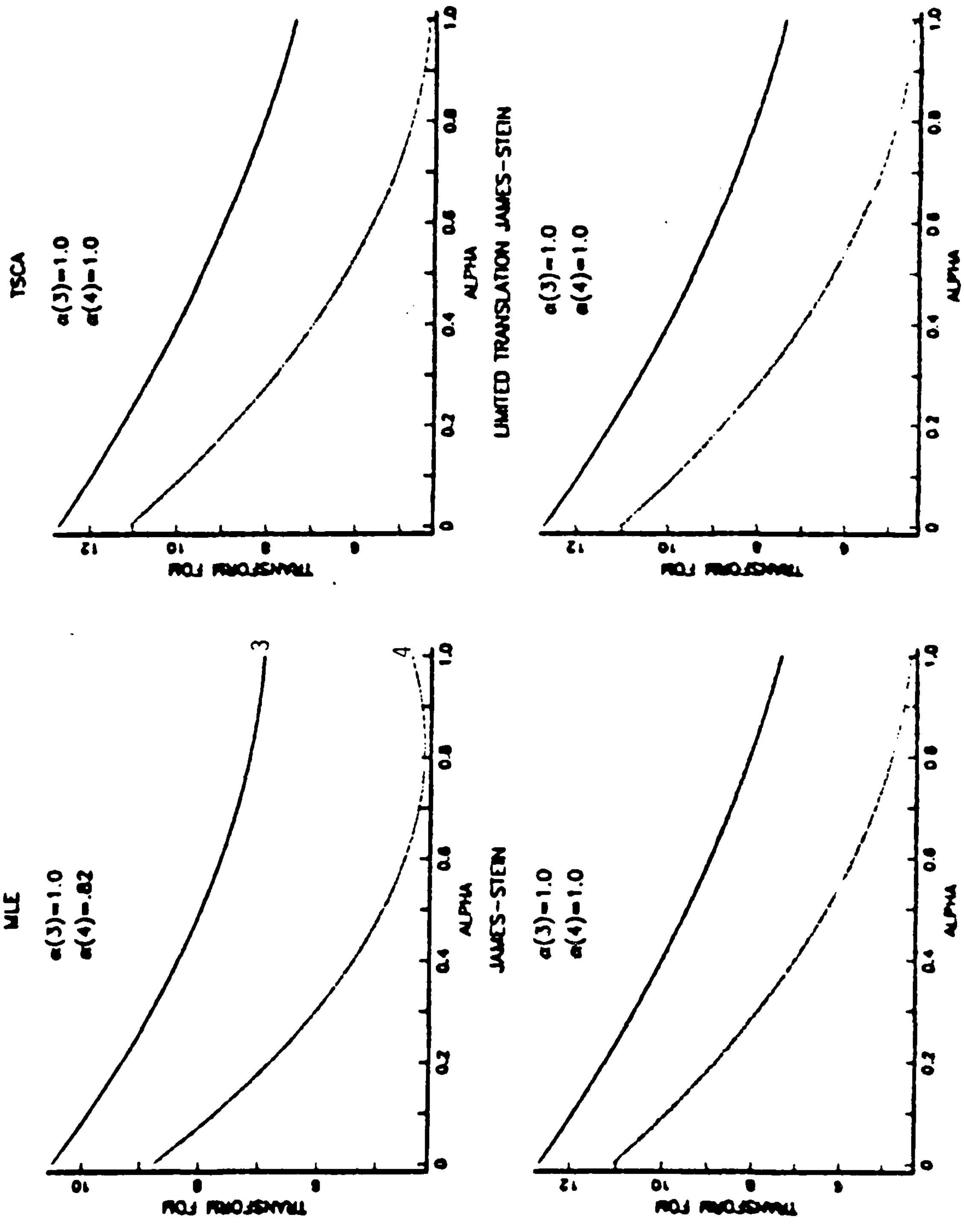

Figure 3.16 Between VY Two Stability -- Aviation LTCOLS

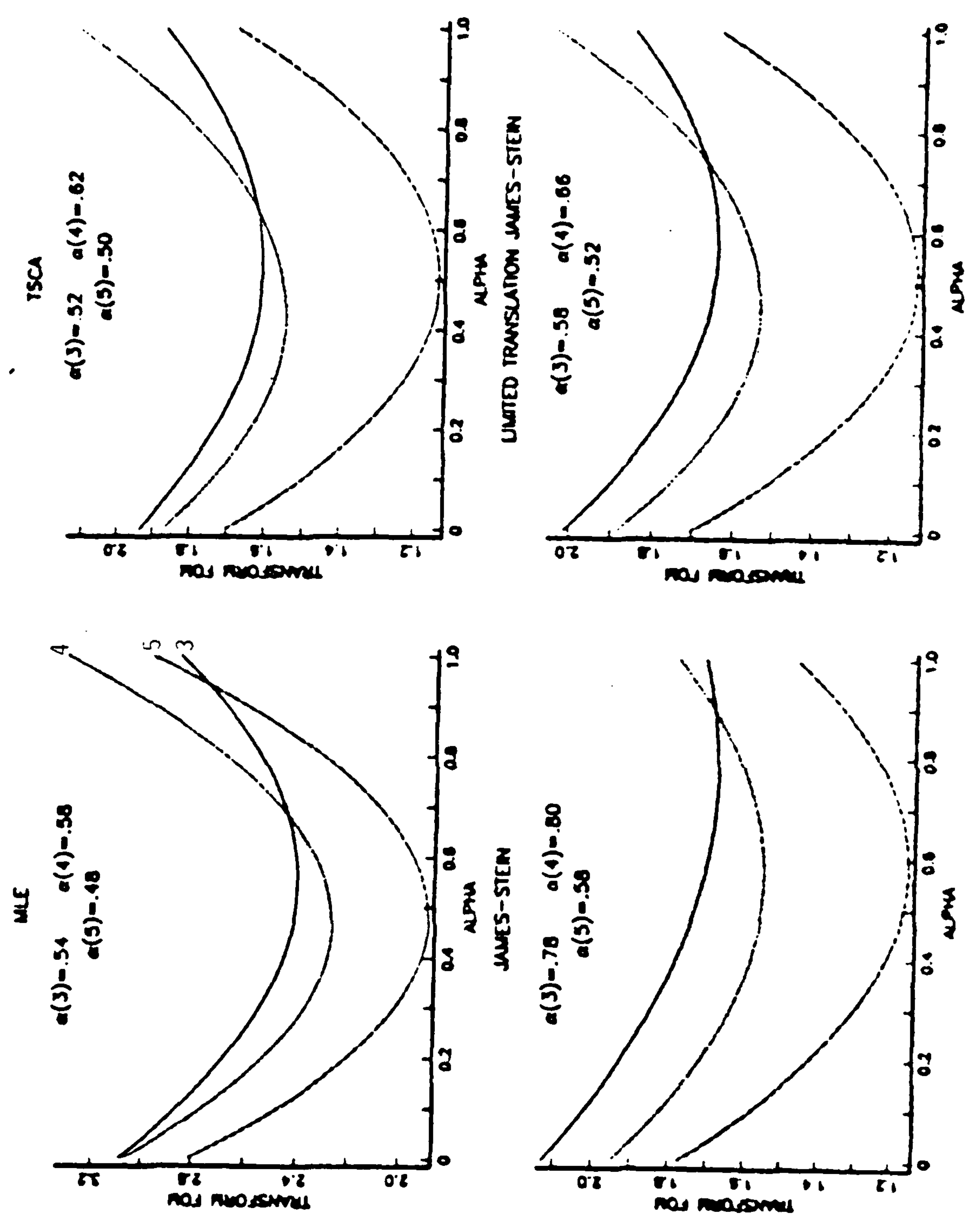

Figure 3.17 Between VY One Stability -- Combat Spt. 1LTS

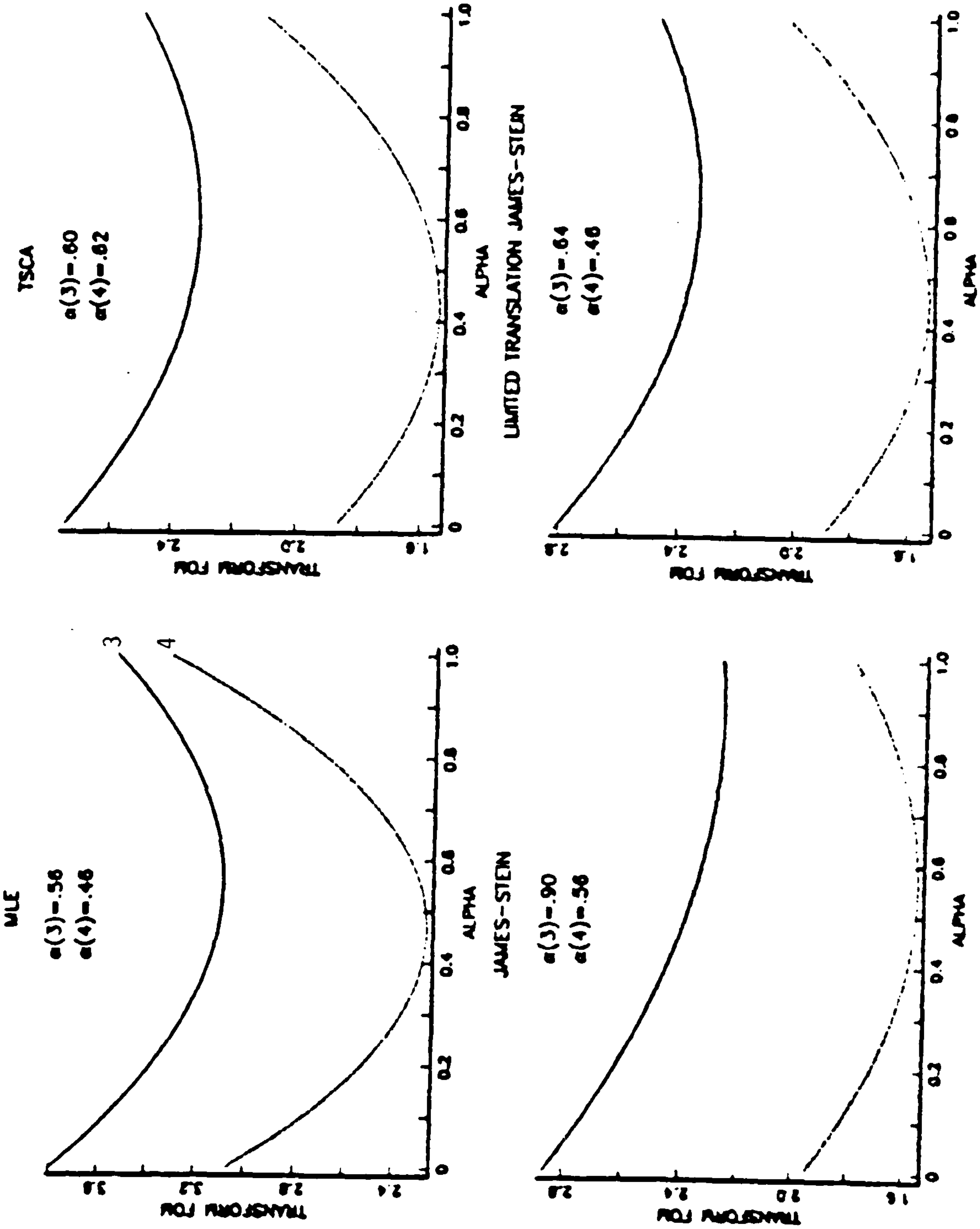

Figure 3.18 Between VY Two Stability -- Combat Spt. 1LTS

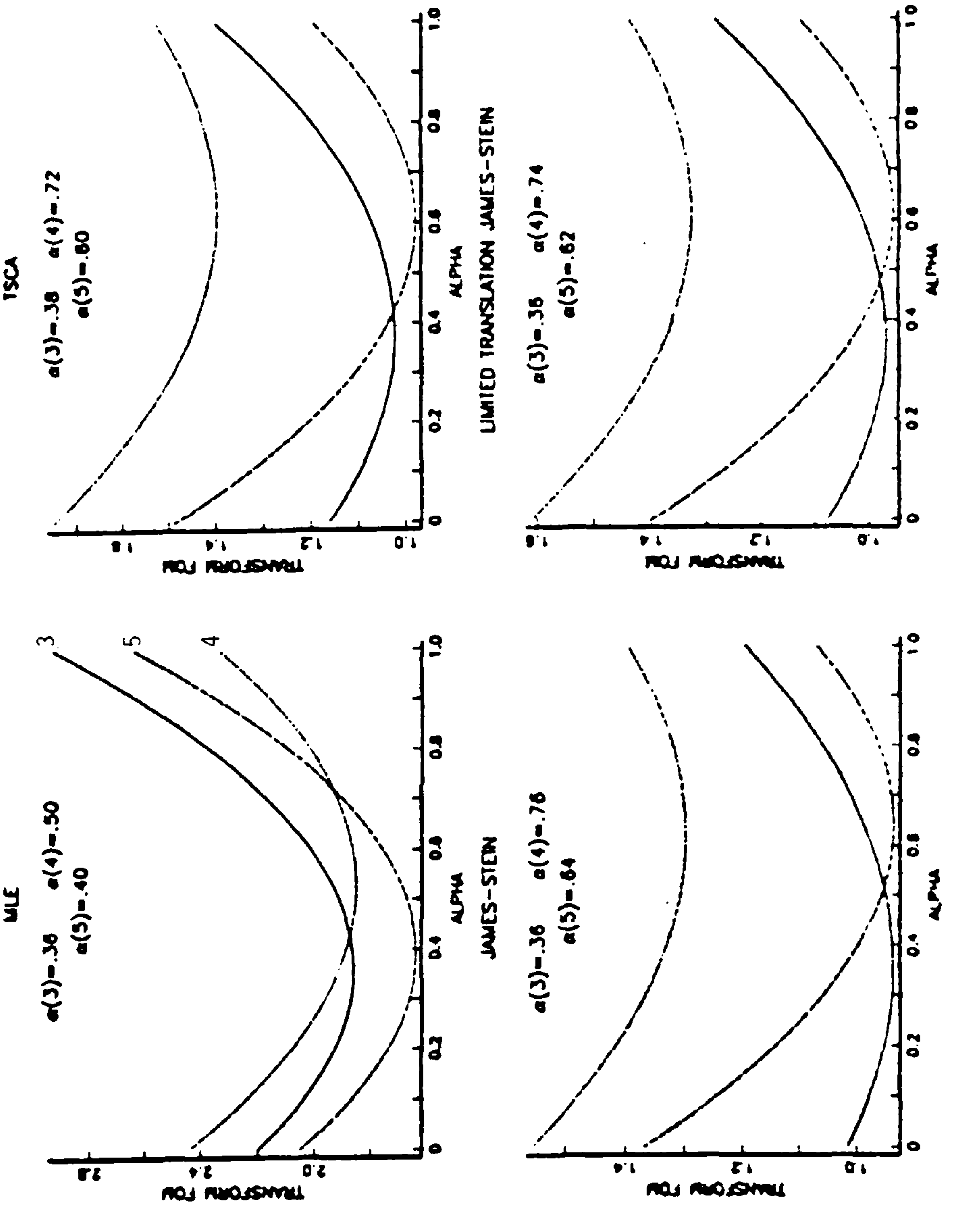

Figure 3.19 Between VY One Stability -- Combat Spt. LTCOLS

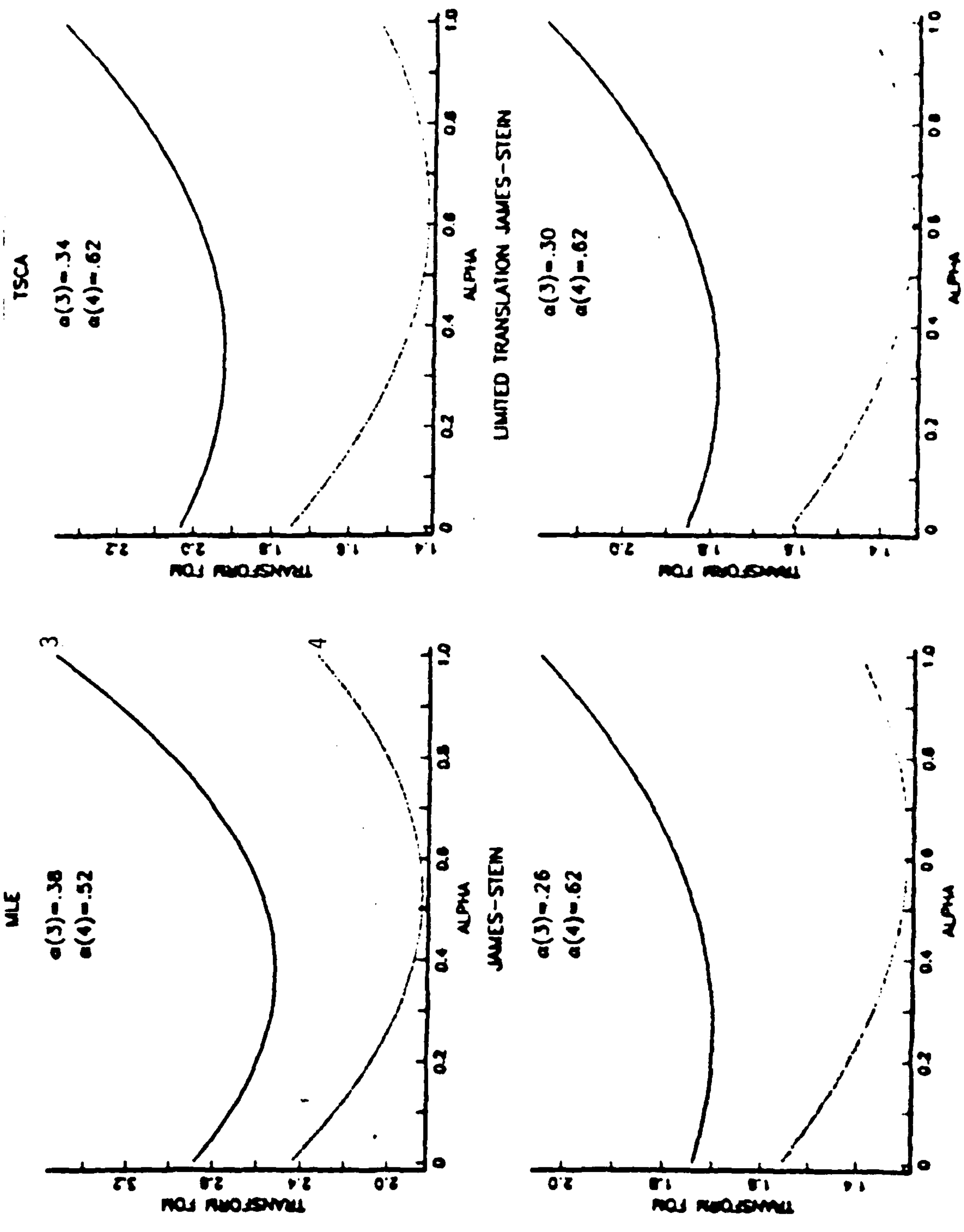

Figure 3.20 Between VY Two Stability -- Combat Spt. LTCOLS

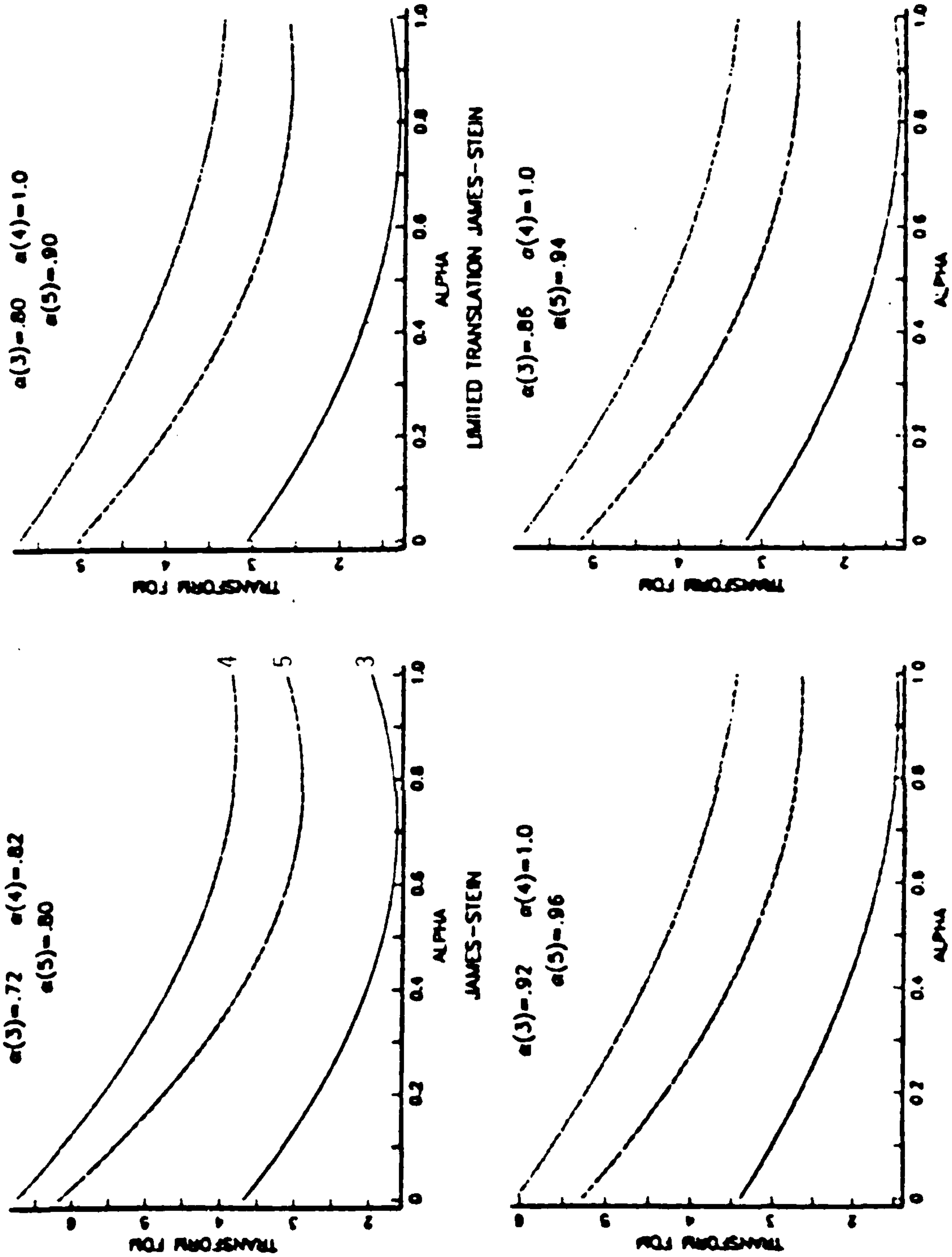

Figure 3.21 Between VY One Stability -- Ground Combat 1LTS

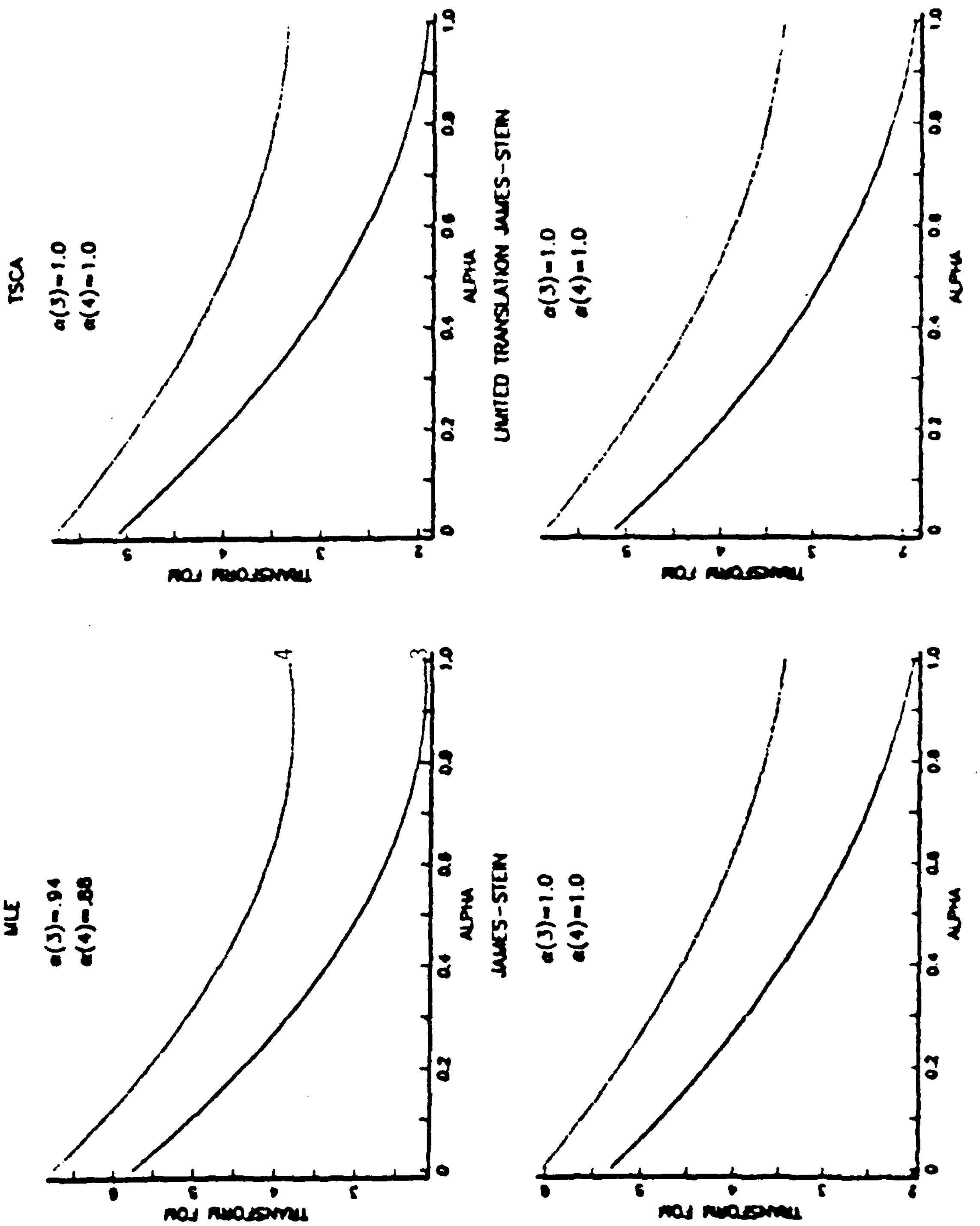

Figure 3.22 Between VY Two Stability -- Ground Combat 1LTS

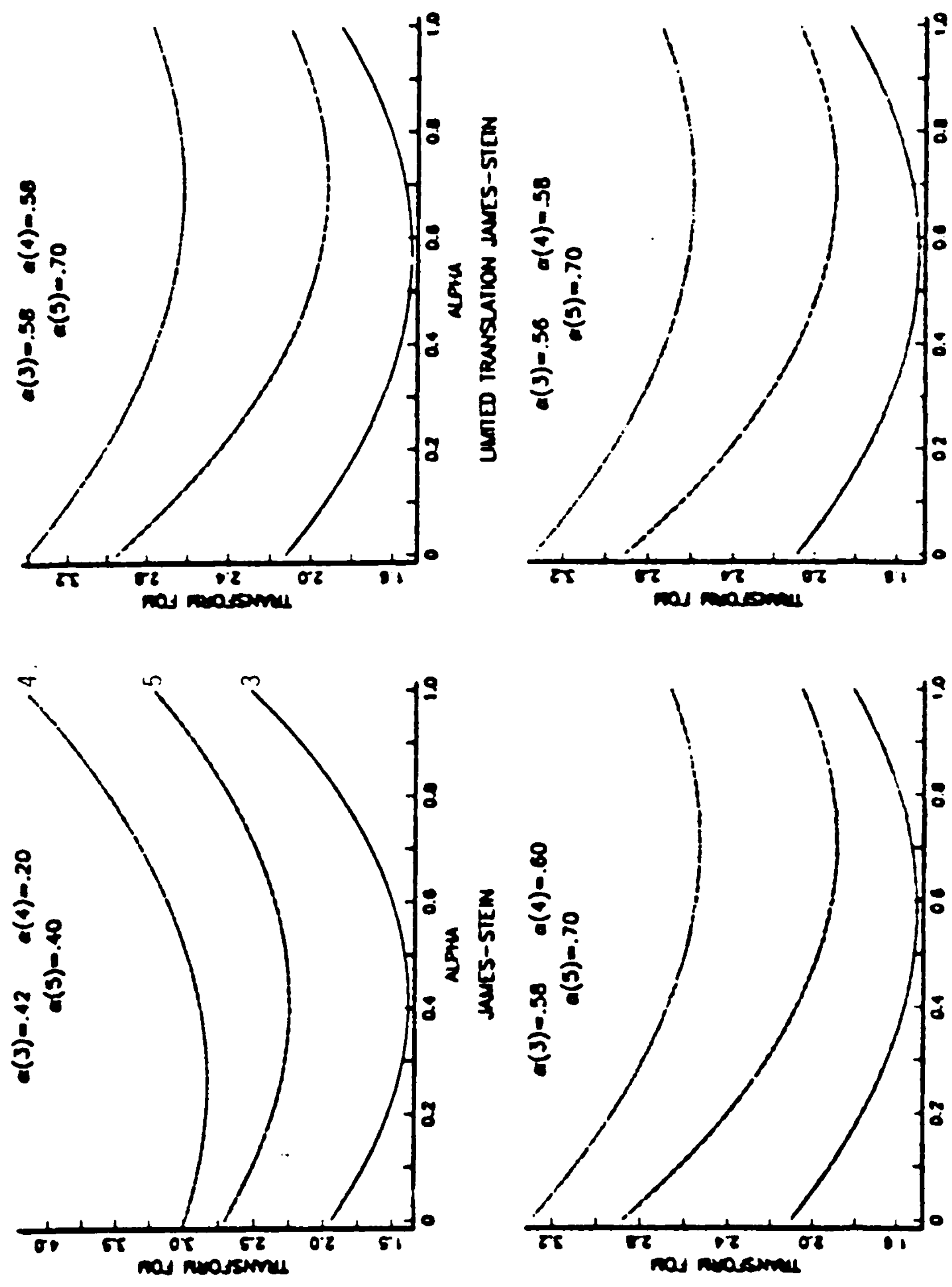

Figure 3.23 Between VY One Stability -- Ground Combat LTCOLS

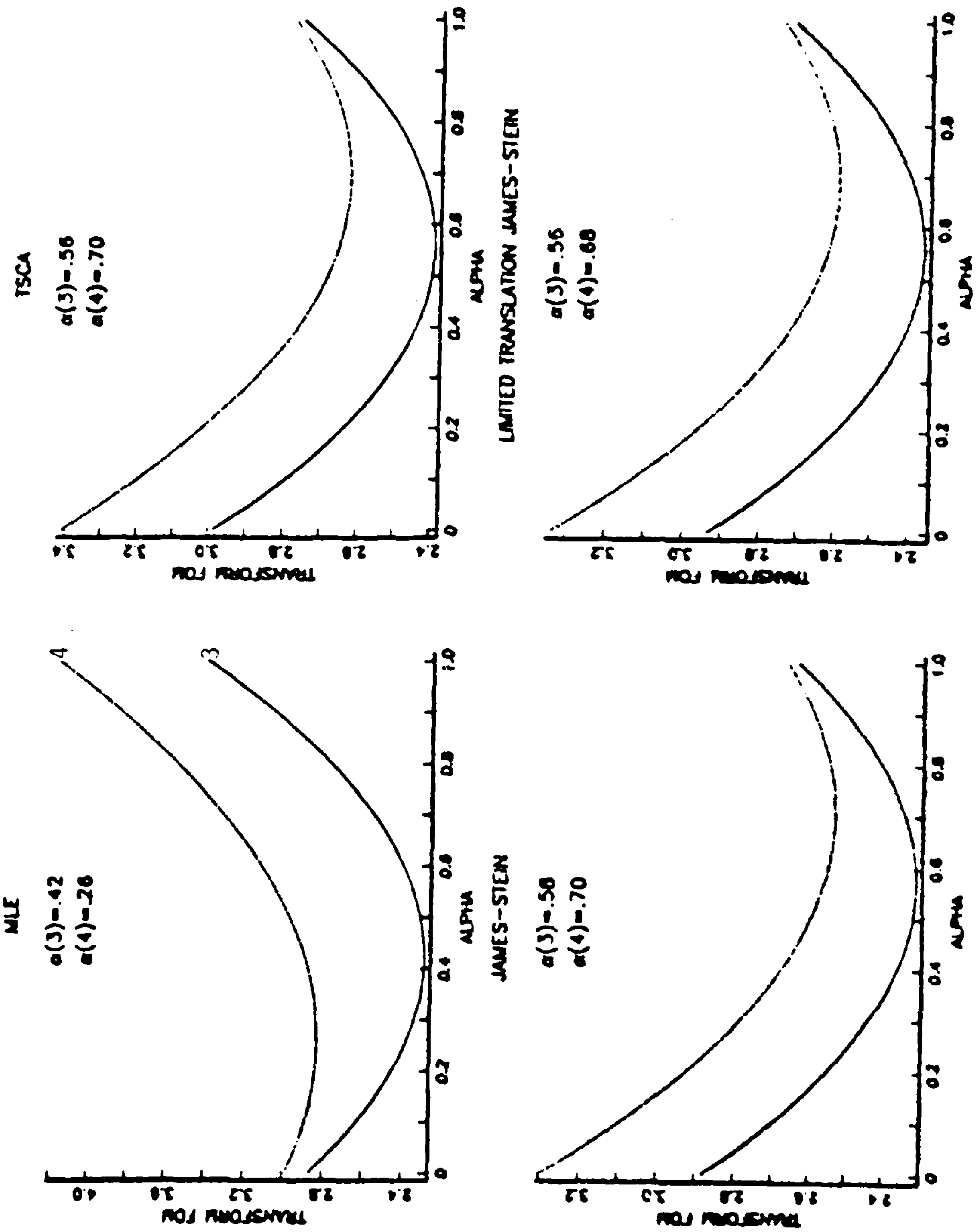

Figure 3.24 Between VY Two Stability -- Ground Combat LTCOLS

line is seen on one of the graphs, usually the 'MLE' graph, on each page.

A review of Figures 3.13-3.24 show mixed results for the stability of the smoothing constants. The aviation aggregate (see Figures 3.13, 3.14, 3.15, and 3.16) shows a large instability between base period lengths for validation year one and an almost perfect stability for validation year two. This is again explained by the initiation of the Aviation Officer Continuation Pay program in 1981. The instability for validation year one is caused by the large decrease in optimal α for the 5-year base period from that of the other two. As discussed in the previous section, the 5-year base contains 1981 data, so the α value no longer has to be 1.0 to produce optimal figures of merit as it did in the 3- and 4-year bases (which explains the stability of α for validation year two). Thus, the instability of α is welcome here, as it is indicating a change in the aviation attrition environment.

From Figures 3.17 and 3.18, we see that α is stable for combat support 1LTS. The only exception is the James-Stein estimator for validation year two. This exception has a percent differential of 10%, while all of the other cases have percent differentials well below the arbitrary acceptable level of 5%. Figures 3.19 and 3.20 show that the opposite is true for combat support LTCOLS; only the MLE shows stability. The percent differentials for TSCA, JS, and LTJS are all above 10% for the validation year one case, and have surprisingly small, though unacceptable, values of 5.1%, 6.7%, and 5.6%, respectively, for the validation year two case. They are "surprisingly" small because the ranges of optimal α are very large, .28, .36, and .32, respectively, and one would think that the percent differential would be much bigger in light of these large ranges. A pairwise analysis of the optimal α values leads us to

believe that the optimal α for the 3-year base is significantly smaller than that for the 4- and 5-year bases, for which α appears stable. Thus, it would seem from this data-limited analysis that the optimal α values for TSCA, JS, and LTJS for combat support LTCOLS is stabilizing as the later years are included in the base.

Figures 3.21 and 3.22 show a very stable optimal α value for all estimators in both validation year cases for ground combat 1LTS, except TSCA for validation year one, which has a percent differential of 8.1%. As mentioned in the previous section, the reasons for the preponderance of α values of 1.0 for this study block are unknown. All that is known is that these large α values indicate a very turbulent environment for the attrition of ground combat 1LTS, with the patterns changing dramatically from year to year. An analysis of Figures 3.23 and 3.24 shows that α is stable for ground combat LTCOLS as well. The only instability seen in these graphs is the MLE for validation year one, which has a percent differential of 6.5%. Therefore, the conclusion is that α is stable for the ground combat aggregate.

IV. RESULTS

A. GENERAL

This chapter displays various data results, using a three-year base period, for the six estimators used in this study.

B. COMPARISON WITH RESULTS FROM EARLIER WORKS

Table 4 thru Table 6 display the figures of merit obtained by using a 3-year base period and smoothing the empirical rate of year 4 (1980) onto it. The figures of merit and optimal α represent those corresponding to the minimum transformed figures of merit for validation year 2 (1982), since it was determined earlier in this study that the ability to forecast two years into the future should be the major concern of the Marine Corps. Table 7 thru Table 9 show the results obtained by Robinson in his thesis [Ref. 2: pp. 39-41]. These estimates are the same as those of a 4-year base period with an α value of 0. In the analysis that follows, this estimation scheme is referred to as the "Robinson method" or the "Robinson estimation scheme."

One may notice a slight difference in the original scale figures of merit between Tables 7, 8, and 9 and Tables 5, 6, and 7 in Robinson's thesis [Ref. 2: pp. 39-41]. This is because a small error was found and corrected in APL function RISKO since the submission of Robinson's thesis in March, 1986 (that being the necessary addition of variable NV on lines 17-18 of the new RISKO, seen in Figure A.8, which was absent from Robinson's version) [Ref. 2]. However, the errors resulting from this mistake in RISKO in Robinson's original scale figures of merit are quite small, and in no way invalidate the comparisons he made [Ref. 2].

The results of using exponential smoothing are now presented for each of the six study blocks. The MOE for comparison is transformed FOM.

TABLE 4

EXPONENTIAL SMOOTHING AVIATION FIGURES OF MERIT

	α	TRANSFORMED FOM 1981	1982	1983
1st LT				
AGG ORIG	1	3.976	5.634	6.260
AGG TRANS	0	206.877	212.181	219.619
MLE	1	2.437	5.746	5.156
TSCA	1	1.243	5.249	4.886
JS	1	1.402	5.360	5.126
LTJS	1	1.386	5.398	5.104
LTCOL				
AGG ORIG	1	7.898	14.829	13.238
AGG TRANS	1	24.733	29.123	29.083
MLE	1	3.423	6.942	7.604
TSCA	1	3.341	7.395	8.210
JS	1	3.302	7.382	8.167
LTJS	1	3.321	7.378	8.191
	α	ORIGINAL FOM 1981	1982	1983
1st LT				
AGG ORIG	1	40.488	12.314	50.363
AGG TRANS	0	384.833	306.151	886.744
MLE	1	13.431	18.361	42.519
TSCA	1	13.428	20.271	25.000
JS	1	16.923	22.791	26.475
LTJS	1	15.797	22.423	26.032
LTCOL				
AGG ORIG	1	74.977	113.067	50.567
AGG TRANS	1	63.264	92.066	76.353
MLE	1	20.801	65.982	32.998
TSCA	1	37.928	44.331	28.742
JS	1	29.665	42.718	27.425
LTJS	1	34.297	43.152	27.835

1. Aviation

From a comparison of Tables 4 and 7, it is seen that exponential smoothing outperforms the Robinson estimation scheme in all cases and for all validation years except the transformed aggregate estimator. This exception is not important at all, since the figures of merit for this

TABLE 5

EXPONENTIAL SMOOTHING COMBAT SUPPORT FIGURES OF MERIT

	α	TRANSFORMED FOM 1981	1982	1983
1st LT				
AGG ORIG	1	1.949	1.868	1.581
AGG TRANS	0	4.086	4.580	3.995
MLE	.56	2.402	3.093	2.420
TSCA	.60	1.617	2.319	1.588
JS	.90	1.688	2.257	1.559
LTJS	.64	1.646	2.336	1.574
LTCOL				
AGG ORIG	0	1.532	1.889	1.956
AGG TRANS	1	3.372	3.649	3.757
MLE	.38	1.860	2.520	2.754
TSCA	.34	1.025	1.923	2.024
JS	.26	.941	1.699	1.732
LTJS	.30	.971	1.782	1.861

	α	ORIGINAL FOM 1981	1982	1983
1st LT				
AGG ORIG	1	88.732	62.779	50.792
AGG TRANS	0	234.581	136.699	96.426
MLE	.56	1262.038	2810.514	3858.857
TSCA	.60	75.051	86.051	65.598
JS	.90	83.608	77.961	47.631
LTJS	.64	75.074	85.040	60.126
LTCOL				
AGG ORIG	0	39.172	38.241	30.186
AGG TRANS	1	41.704	38.615	33.363
MLE	.38	136.467	38.763	157.663
TSCA	.34	23.446	31.130	29.101
JS	.26	22.863	29.503	26.596
LTJS	.30	23.074	29.367	27.342

estimator in all cases are so much higher than those for the others. This result extends into original scale, where exponential smoothing is again seen to have lower figures of merit. Special notice should be given to the exceptional performance of exponential smoothing in forecasting rates two and three years into the future as compared to the Robinson method.

A major reason that exponential smoothing does so well in this case is the inflexibility of the Robinson estimation scheme. His scheme cannot readily adjust to changes

TABLE 6

EXPONENTIAL SMOOTHING GROUND COMBAT FIGURES OF MERIT

		TRANSFORMED FOM		
	α	1981	1982	1983
1st LT				
AGG ORIG	.92	3.045	3.663	5.988
AGG TRANS	0	19.842	20.724	24.011
MLE	.94	1.858	2.165	4.555
TSCA	1	1.431	1.916	4.424
JS	1	1.472	2.008	4.566
LTJS	1	1.405	1.869	4.415
LTCOL				
AGG ORIG	1	3.540	3.427	3.671
AGG TRANS	0	14.393	13.453	13.777
MLE	.42	1.389	2.293	3.085
TSCA	.56	1.510	2.392	3.169
JS	.58	1.505	2.313	3.039
LTJS	.56	1.494	2.358	3.108

		ORIGINAL FOM		
	α	1981	1982	1983
1st LT				
AGG ORIG	.92	64.779	62.206	80.907
AGG TRANS	0	235.108	236.736	303.748
MLE	.94	175558.747	449.816	3491.851
TSCA	1	75.514	57.155	124.972
JS	1	82.814	55.067	122.684
LTJS	1	78.815	53.724	128.000
LTCOL				
AGG ORIG	1	90.948	110.793	69.991
AGG TRANS	0	221.432	197.458	128.373
MLE	.42	37.820	1383.609	38.934
TSCA	.56	40.713	52.936	49.449
JS	.58	40.615	53.345	48.090
LTJS	.56	40.111	52.986	48.151

in the environment such as the Aviation Officer Continuation Pay program initiated in 1981. This program, as has been mentioned before, is believed to have radically changed attrition patterns for aviators as compared to 1977-1980. The exponential smoothing model's ability to anticipate this change, which it does by completely eliminating the effects of data from years 1977-1979 by having α values of 1, allows it to better predict attrition rates for the years following the change in the aviation environment. Eventually, once a base period of post-1981 data is established, the α values

TABLE 7

FOUR-YEAR BASE ESTIMATE AVIATION FIGURES OF MERIT

	TRANSFORMED FOM		
	1981	1982	1983
1st LT			
AGG ORIG	6.405	9.645	11.393
AGG TRANS	197.621	202.022	208.808
MLE	3.914	9.981	10.420
TSCA	3.461	9.574	10.042
JS	3.678	9.768	10.318
LTJS	3.642	9.764	10.279
LTCOL			
AGG ORIG	9.957	17.997	15.488
AGG TRANS	25.093	29.506	29.310
MLE	4.366	9.058	8.210
TSCA	5.777	10.967	10.394
JS	5.737	10.911	10.355
LTJS	5.744	10.901	10.340
	ORIGINAL FOM		
	1981	1982	1983
1st LT			
AGG ORIG	31.499	34.924	54.911
AGG TRANS	321.894	286.068	572.830
MLE	22.804	46.751	57.475
TSCA	19.333	38.458	49.389
JS	21.144	39.596	51.013
LTJS	20.732	39.434	50.690
LTCOL			
AGG ORIG	74.697	111.913	51.256
AGG TRANS	23050.668	248.171	3900.284
MLE	33.932	55.925	22.116
TSCA	37.974	57.475	37.637
JS	36.708	54.901	35.218
LTJS	37.180	55.414	35.619

should decrease substantially, thereby giving the base estimate some meaning in the estimation of the attrition rates. Further study into this matter should be undertaken as more data becomes available.

2. Combat Support

From a comparison of tables 5 and 8, one sees that in all cases for combat support 1LTS, the Robinson method figures of merit are slightly smaller than those produced by exponential smoothing. The values of α vary widely over the six estimators. Notice that maximum likelihood estimation

TABLE 8

FOUR-YEAR BASE ESTIMATE COMBAT SUPPORT FIGURES OF MERIT

	TRANSFORMED FOM		
	1981	1982	1983
1st LT			
AGG ORIG	1.528	1.842	1.385
AGG TRANS	3.308	3.722	3.238
MLE	2.273	2.843	2.409
TSCA	1.383	2.008	1.468
JS	1.470	2.066	1.532
LTJS	1.428	2.045	1.482
LTCOL			
AGG ORIG	1.316	1.701	1.755
AGG TRANS	2.601	2.898	2.981
MLE	1.589	2.157	2.444
TSCA	.910	1.689	1.752
JS	.809	1.514	1.556
LTJS	.831	1.560	1.611
	ORIGINAL FOM		
	1981	1982	1983
1st LT			
AGG ORIG	79.521	69.344	54.319
AGG TRANS	288.130	131.044	104.613
MLE	141.815	145.434	104.668
TSCA	73.190	85.451	71.852
JS	74.158	81.133	63.562
LTJS	74.148	84.524	68.079
LTCOL			
AGG ORIG	34.631	46.427	38.481
AGG TRANS	41.785	44.739	64.345
MLE	36.015	62.882	48.468
TSCA	27.773	40.777	45.453
JS	26.191	36.758	41.956
LTJS	26.171	36.865	42.044

produces alarmingly high figures of merit in the original scale. This is a direct result of the unsuitability of using the optimal α for transformed scale to predict original scale rates discussed in Chapter 3 (see also Appendix C). Additionally, its figures of merit in transformed scale are higher than those for all but one of the other estimators, so it appears that MLE is not a very good alternative to use for this study block.

TABLE 9

FOUR-YEAR BASE ESTIMATE GROUND COMBAT FIGURES OF MERIT

	TRANSFORMED FOM		
	1981	1982	1983
1st LT			
AGG ORIG	2.774	3.609	5.720
AGG TRANS	18.280	19.064	22.044
MLE	2.693	4.258	6.178
TSCA	1.957	3.447	5.231
JS	2.169	3.629	5.505
LTJS	2.045	3.482	5.334
LTCOL			
AGG ORIG	3.692	3.596	3.783
AGG TRANS	13.108	12.209	12.546
MLE	1.252	2.029	2.925
TSCA	1.598	2.414	3.333
JS	1.584	2.332	3.223
LTJS	1.567	2.367	3.263
	ORIGINAL FOM		
	1981	1982	1983
1st LT			
AGG ORIG	79.406	87.445	110.324
AGG TRANS	271.834	255.215	341.258
MLE	91.399	134.172	170.424
TSCA	69.166	94.099	118.799
JS	76.811	98.672	125.292
LTJS	73.141	94.657	121.432
LTCOL			
AGG ORIG	100.321	122.612	75.718
AGG TRANS	250.966	342.375	4494.138
MLE	35.833	45.817	56.710
TSCA	37.521	51.760	57.158
JS	36.909	50.712	55.615
LTJS	36.373	50.828	55.562

Comparing Tables 5 and 8 for lieutenant colonels, we see that the Robinson method's estimates are barely better than the exponential smoothing estimates in all cases. The values of α producing the minimum figures of merit for 1982 range from .26 to .38 over the 4 estimators introduced by Robinson [Ref. 2]. However, as one can see from Figure 3.4, the curves are flat enough in this range that choosing any α in this range would produce good results.

Exponential smoothing, therefore, proves to be an good technique for predicting attrition rates for the combat

support aggregate. Its forecasting ability is almost as good as Robinson's method while requiring far less data.

3. Ground Combat

A comparison of Tables 6 and 9 shows that exponential smoothing outperforms the Robinson method's estimates in most cases for the ground combat aggregate. For 1LTS, exponential smoothing is better in all cases except for the aggregate estimators. The improvement over the Robinson method figures of merit is almost 50% for validation year 1982 for MLE, TSCA, JS, and LTJS, and is also quite noticeable for the other two validation years. Again, the reasons for the α values near or at 1.0 are unknown, but they indicate a radical change in the attrition environment for ground combat 1LTS. Further study is needed to determine whether or not this turbulence is specific to our data. But whatever causes this apparent yearly change in attrition patterns, the exponential smoothing model shows its efficacy as a forecasting model in this study block by anticipating this change and discounting the now-irrelevant base period data, with the result being significant improvements in forecasting ability over the methods of Robinson.

For ground combat LTCOLS, The AGG ORIG, TSCA, JS, and LTJS estimators using exponential smoothing have smaller figures of merit than the Robinson estimation scheme, and the figures of merit for AGG TRANS and MLE are slightly higher for exponential smoothing than for the Robinson method. However, the differences are small in both directions; they are not nearly as pronounced for this study block as they are for combat support 1LTS. These results hold for all three validation years. Again, we see the exponential smoothing model producing attrition rates as good as, if not better than, a method which theoretically requires much more data. This study block, then, also shows exponential smoothing to be a good model to perform the attrition rate calculations.

C. ATTRITION RATES

The attrition rates for all 7 operational fields for first lieutenants and lieutenant colonels are presented in Tables 10 thru 23. Attrition rates are computed for all 6 estimators and all 31 lengths of service.

Unlike the estimates produced by Robinson for the original aggregate [Ref. 2: pp. 54-67], those calculated by exponential smoothing are not the same for all cells with non-zero rates, except when $\alpha=0$, as is the case with combat support LTCOLS. They would be the same were the smoothing done in original scale, but since it is done in transformed scale, variability is introduced. This is because taking the average of two sets of data in transformed scale, linearly combining them using Equation 2.1, and then transforming back to original scale will not produce a single average for all cells as would linearly combining two original scale average rates. This is basically the same reason that there is variability in the transformed aggregate estimates.

TABLE 10

AVIATION ATTRITION RATES FOR 1ST LTS

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	.0179	.1154	.0008	.0003	.0005	.0005
3	.0253	.2283	.0266	.0283	.0300	.0294
4	.0159	.2527	.0090	.0100	.0114	.0106
5	.0282	.2221	.0402	.0421	.0437	.0434
6	.0216	.1442	.0240	.0270	.0278	.0278
7	.0204	.0758	.0386	.0432	.0430	.0430
8	.0350	.0064	.0621	.0729	.0698	.0698
9	.0788	.0067	.0251	.0533	.0510	.0510
10	.0710	.0072	.1197	.1335	.1282	.1282
11	.0480	.0182	.0063	.0307	.0277	.0277
12	.0123	0	.0033	.0155	.0097	.0109
13	.0832	0	.0284	.2679	.2027	.2313
14	.0725	0	.5000	.5000	.5000	.5000
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	0	0	0	0	0	0
19	0	0	0	0	0	0
20	0	0	0	0	0	0
21	0	0	0	0	0	0
22	0	0	0	0	0	0
23	0	0	0	0	0	0
24	0	0	0	0	0	0
25	0	0	0	0	0	0
26	0	0	0	0	0	0
27	0	0	0	0	0	0
28	0	0	0	0	0	0
29	0	0	0	0	0	0
30	0	0	0	0	0	0

TABLE 11

AVIATION ATTRITION RATES FOR LTCOLS

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	.0098	0	.2957	.0298	.0373	.0373
16	.0089	.1001	.0001	.0101	.0076	.0076
17	.0907	.0672	.0270	.0010	.0003	.0004
18	.1304	.1762	.0015	.0026	.0046	.0034
19	.1302	.2278	.0015	.0012	.0029	.0016
20	.1442	.2514	.0717	.0790	.0850	.0820
21	.1739	.2575	.1568	.1642	.1684	.1681
22	.1791	.2639	.1991	.2054	.2084	.2084
23	.1812	.2512	.2447	.2506	.2514	.2514
24	.1919	.2182	.2464	.2549	.2542	.2542
25	.1776	.1835	.2676	.2765	.2735	.2735
26	.1281	.1071	.3326	.3401	.3303	.3313
27	.1018	.0237	.2925	.3070	.2926	.2950
28	.1602	.0016	.6575	.6311	.5953	.6145
29	.1789	.0430	.4119	.4343	.3998	.4139
30	.1184	.0500	.9944	.9683	.9423	.9605

TABLE 12

COMBAT SUPPORT ATTRITION RATES FOR 1ST LTS
CODE 07 ENGINEERS

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	.2357	.3571	.0901	.1119	.1388	.1134
3	.1987	.3973	.2610	.2697	.3124	.2791
4	.1547	.3940	.1446	.1455	.1768	.1454
5	.1903	.3793	.1838	.1861	.2144	.1897
6	.1870	.3511	.2990	.3072	.3086	.3135
7	.2066	.2976	.3226	.3443	.3214	.3325
8	.1820	.1756	.3753	.4145	.2619	.3768
9	.1824	.1916	.0520	.1312	.0871	.1266
10	.1933	.2023	.2075	.2523	.2371	.2444
11	.2578	.1953	.2667	.3089	.3358	.3200
12	.2370	.1711	.0411	.1327	.1338	.1326
13	.1915	.1461	0	.0801	.0843	.0868
14	.1216	.1711	.0064	.0269	.0273	.0329
15	.1375	.1618	.0036	.0373	.0384	.0463
16	.1683	.1218	.0005	.0748	.0640	.0754
17	.1616	.1149	.0009	.0739	.0586	.0723
18	.2148	.1461	.0012	.0966	.1065	.0977
19	.2209	.1218	.0018	.1161	.1132	.1113
20	.2370	.1002	.0646	.1989	.1384	.1692
21	.2869	.1077	.0924	.2503	.1946	.2252
22	.2756	.0923	.0118	.1912	.1728	.1845
23	.2062	.0661	.0006	.1463	.1005	.1220
24	.2815	.0168	.1583	.3626	.4550	.3571
25	.2370	.0003	.1671	.4058	.4575	.3916
26	.2370	.0003	.1671	.4058	.4575	.3916
27	0	0	0	0	0	0
28	0	0	0	0	0	0
29	0	0	0	0	0	0
30	0	0	0	0	0	0

TABLE 13

COMBAT SUPPORT ATTRITION RATES FOR 1ST LTS
CODE 13 COMMUNICATIONS

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	.2370	.3579	.1800	.1738	.2067	.1802
3	.1939	.4065	.2699	.2669	.2949	.2719
4	.1855	.4031	.2084	.2063	.2264	.2063
5	.2148	.3893	.3155	.3184	.3204	.3245
6	.2086	.3540	.2932	.3072	.3356	.3087
7	.2077	.2932	.4424	.4523	.3931	.4370
8	.2561	.2847	.2528	.2799	.2783	.2767
9	.2291	.3230	.1881	.2078	.2178	.2172
10	.2091	.3304	.1381	.1514	.2213	.1667
11	.1977	.2909	.0941	.1268	.1807	.1422
12	.1715	.2179	.0097	.0579	.0681	.0663
13	.1332	.1077	.0175	.0942	.0410	.0857
14	.1350	.0367	.0040	.1113	.0398	.0774
15	.0905	.0003	.0265	.1909	.0222	.1426
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	0	0	0	0	0	0
19	.2370	.0003	.1671	.4058	.4575	.3916
20	.2370	.0003	.1671	.4058	.4575	.3916
21	0	0	0	0	0	0
22	0	0	0	0	0	0
23	0	0	0	0	0	0
24	0	0	0	0	0	0
25	0	0	0	0	0	0
26	0	0	0	0	0	0
27	0	0	0	0	0	0
28	0	0	0	0	0	0
29	0	0	0	0	0	0
30	0	0	0	0	0	0

TABLE 14

COMBAT SUPPORT ATTRITION RATES FOR 1ST LTS CODE 20 MOTOR TRANSPORT

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	.0905	.1618	.0140	.0173	.0102	.0172
3	.0983	.2472	.2572	.2942	.2035	.2776
4	.0926	.2690	.2669	.2635	.2461	.2534
5	.0812	.2751	.1576	.2076	.2071	.2178
6	.0670	.2149	.2195	.2537	.1756	.2345
7	.0817	.0752	.1060	.2807	.1599	.2411
8	.1616	.0266	.9991	.8392	.5214	.8041
9	.1271	.0566	.9236	.8146	.6270	.7987
10	.1746	.0367	.0002	.1614	.0738	.1243
11	.1721	.1002	.1863	.3277	.2873	.3184
12	.1721	.1002	.0003	.0938	.0684	.0805
13	.1528	.1461	.0017	.0544	.0508	.0652
14	.1862	.1404	0	.0839	.0803	.0881
15	.2692	.1798	.0102	.1228	.1626	.1249
16	.2043	.1345	.0005	.0949	.0966	.0950
17	.1862	.0840	0	.1137	.0811	.0938
18	.2546	.0661	.0070	.1956	.1508	.1759
19	.2148	.0752	.0012	.1508	.1094	.1288
20	.3091	.0467	.3740	.4019	.4646	.3950
21	.2546	.0661	.0070	.1956	.1508	.1759
22	.3011	.0367	.1552	.3486	.4546	.3463
23	.3011	.0367	.1552	.3233	.4502	.3243
24	.2692	.0081	.1605	.3767	.4561	.3686
25	.2692	.0081	.1605	.3731	.4555	.3654
26	.2692	.0081	.1605	.3731	.4555	.3654
27	.1683	.0018	.0005	.2301	.0703	.1769
28	.1216	.0088	.0516	.2748	.0477	.2210
29	.2148	.0088	.1726	.4482	.4620	.4272
30	.2692	.0081	.1605	.3559	.4525	.3500

TABLE 15

COMBAT SUPPORT ATTRITION RATES FOR LTCOLS
CODE 07 ENGINEERS

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	.2019	.0094	.1263	.3156	.3495	.3273
16	0	0	0	0	0	0
17	.1404	.1507	.0001	.1012	.1084	.1100
18	.1531	.2705	.0001	.0370	.0780	.0546
19	.1542	.2939	.0001	.0299	.0739	.0443
20	.1422	.3012	.2587	.2506	.2557	.2589
21	.1328	.2984	.0443	.0897	.1303	.1100
22	.1408	.2908	.1488	.1486	.1722	.1660
23	.1514	.2662	.0962	.1260	.1492	.1452
24	.1366	.2368	.2638	.2742	.2543	.2591
25	.1492	.1979	.1442	.2581	.2640	.2534
26	.1746	.0977	.0007	.1273	.1289	.1237
27	.1475	.0866	.6841	.6243	.5213	.5931
28	.1289	.0218	.6117	.5407	.4169	.4692
29	.1289	.0218	.0685	.3187	.2386	.2514
30	.1206	.0748	.5488	.4937	.4012	.4345

TABLE 16

COMBAT SUPPORT ATTRITION RATES FOR LTCOLS
CODE 13 COMMUNICATIONS

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	0	0	0	0	0	0
16	.1385	.0010	.0777	.4148	.2942	.3253
17	.1200	.1767	.0052	.1298	.1252	.1324
18	.1304	.2615	.0012	.0530	.0831	.0681
19	.1535	.3052	.0001	.0273	.0723	.0403
20	.1381	.3222	.1480	.1784	.2087	.2010
21	.1238	.3170	.1689	.1914	.2051	.1967
22	.1515	.3170	.1590	.1927	.2193	.2170
23	.1673	.2841	.1437	.1654	.1799	.1763
24	.1495	.2615	.2642	.2531	.2573	.2559
25	.1488	.2302	.1832	.2521	.2699	.2524
26	.1369	.1644	.3054	.3026	.2411	.2668
27	.1637	.1177	.3865	.4655	.3164	.3906
28	.1475	.0866	0	.1736	.1510	.1499
29	.1475	.0866	.1387	.3213	.2692	.2749
30	.1321	.2368	.3756	.3945	.3490	.3531

TABLE 17

COMBAT SUPPORT ATTRITION RATES FOR LTCOLS
CODE 20 MOTOR TRANSPORT

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0						
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	0	0	0	0	0	0
16	.1780	.0010	.1573	.3692	.3769	.3584
17	.1780	.0010	.1573	.3692	.3769	.3584
18	.1637	.0094	.0003	.3062	.2117	.2227
19	.1475	.0354	0	.2113	.1550	.1551
20	.1385	.0623	.1923	.3483	.2709	.2801
21	.1238	.0491	.4323	.4363	.3445	.3686
22	.1424	.0491	.2332	.4020	.3082	.3311
23	.1780	.0010	.1573	.3692	.3769	.3584
24	0	0	0	0	0	0
25	0	0	0	0	0	0
26	0	0	0	0	0	0
27	0	0	0	0	0	0
28	0	0	0	0	0	0
29	.1637	.0094	.0003	.3062	.2117	.2227
30	.1570	.0623	.8041	.6442	.4519	.5405

TABLE 18

GROUND COMBAT ATTRITION RATES FOR 1ST LTS
CODE 03 INFANTRY

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0	0	0	0	0	0	0
1	.1835	0	.4514	.5000	.5000	.5000
2	.1861	.3367	.0129	.0217	.0385	.0246
3	.1690	.3828	.1093	.1063	.1291	.1106
4	.1449	.3829	.1458	.1423	.1632	.1472
5	.1808	.3725	.1732	.1764	.1937	.1822
6	.1792	.3453	.2387	.2407	.2486	.2486
7	.1474	.2996	.2559	.2637	.2630	.2630
8	.1859	.2312	.0855	.1026	.1105	.1105
9	.2204	.2653	.2155	.2353	.2333	.2333
10	.1974	.2649	.1681	.1832	.1870	.1870
11	.1847	.2271	.0814	.1031	.1105	.1105
12	.1927	.1047	.1769	.2217	.1971	.2000
13	.1722	.0208	.0038	.0694	.0547	.0547
14	.1913	.0566	.0094	.1855	.1114	.1414
15	.2367	.0566	.0570	.2892	.1882	.2366
16	.1949	.1904	.0113	.2494	.1371	.1898
17	0	0	0	0	0	0
18	0	0	0	0	0	0
19	0	0	0	0	0	0
20	0	0	0	0	0	0
21	0	0	0	0	0	0
22	0	0	0	0	0	0
23	0	0	0	0	0	0
24	0	0	0	0	0	0
25	0	0	0	0	0	0
26	0	0	0	0	0	0
27	0	0	0	0	0	0
28	0	0	0	0	0	0
29	0	0	0	0	0	0
30	0	0	0	0	0	0

TABLE 19

GROUND COMBAT ATTRITION RATES FOR 1ST LTS
CODE 05 ARTILLERY

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	.1879	.2913	.0126	.0285	.0435	.0328
3	.1578	.3427	.1780	.1797	.1932	.1869
4	.1520	.3442	.2301	.2318	.2404	.2396
5	.1682	.3218	.1889	.1938	.2035	.2022
6	.1597	.2750	.2756	.2793	.2734	.2734
7	.1467	.1705	.4580	.4711	.4242	.4502
8	.1707	.1329	.1988	.2302	.2093	.2103
9	.1684	.1424	.2686	.2891	.2614	.2682
10	.1889	.1537	.1291	.1644	.1559	.1559
11	.1741	.0922	.1440	.1918	.1695	.1703
12	.1364	.0369	.1185	.1753	.1441	.1498
13	.0906	.0027	.1420	.2080	.1458	.1698
14	.1328	.2399	.0047	.1210	.0486	.0761
15	.2050	0	.4504	.5000	.5000	.5000
16	.2050	0	.4504	.5000	.5000	.5000
17	0	0	0	0	0	0
18	0	0	0	0	0	0
19	0	0	0	0	0	0
20	0	0	0	0	0	0
21	0	0	0	0	0	0
22	0	0	0	0	0	0
23	0	0	0	0	0	0
24	0	0	0	0	0	0
25	0	0	0	0	0	0
26	0	0	0	0	0	0
27	0	0	0	0	0	0
28	0	0	0	0	0	0
29	0	0	0	0	0	0
30	0	0	0	0	0	0

TABLE 20

GROUND COMBAT ATTRITION RATES FOR 1ST LTS
CODE 10 TANKS

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	.2003	.1607	.0256	.0654	.0694	.0694
3	.1756	.2534	.1063	.1121	.1215	.1215
4	.1287	.2543	.1321	.1326	.1403	.1403
5	.1807	.2516	.2540	.2716	.2633	.2633
6	.1663	.2019	.2822	.2899	.2719	.2731
7	.1307	.1065	.0716	.0850	.0808	.0808
8	.1471	.0208	0	.0355	.0279	.0279
9	.2572	.0392	.0830	.2134	.1760	.1861
10	.2088	.0683	.0271	.1059	.0932	.0932
11	.1391	.0230	.0011	.0258	.0206	.0206
12	.1479	.0731	.0006	.1037	.0526	.0691
13	.1192	0	.0124	.1062	.0364	.0622
14	.0958	0	.0219	.1489	.0331	.0813
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	.2050	0	.4504	.5000	.5000	.5000
18	0	0	0	0	0	0
19	0	0	0	0	0	0
20	0	0	0	0	0	0
21	0	0	0	0	0	0
22	0	0	0	0	0	0
23	0	0	0	0	0	0
24	0	0	0	0	0	0
25	0	0	0	0	0	0
26	0	0	0	0	0	0
27	0	0	0	0	0	0
28	0	0	0	0	0	0
29	0	0	0	0	0	0
30	0	0	0	0	0	0

TABLE 21

GROUND COMBAT ATTRITION RATES FOR LTCOLS
CODE 03 INFANTRY

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	.0151	.1625	.0596	.1020	.0867	.0964
16	.0277	.0191	.0034	.0150	.0122	.0151
17	.1500	.2018	.0001	.0204	.0246	.0226
18	.1327	.3044	0	.0054	.0089	.0061
19	.1195	.3290	.0029	.0058	.0092	.0064
20	.1311	.3383	.1187	.1211	.1281	.1237
21	.1476	.3345	.0848	.1002	.1091	.1027
22	.1411	.3277	.1395	.1592	.1673	.1623
23	.1314	.3099	.1477	.1667	.1736	.1702
24	.1627	.2941	.2292	.2490	.2528	.2508
25	.1466	.2590	.2207	.2388	.2410	.2396
26	.1394	.2159	.2107	.2034	.2009	.2039
27	.1345	.1618	.2265	.2134	.2062	.2114
28	.1707	.0897	.2299	.2495	.2424	.2415
29	.1741	.0661	.1263	.1549	.1471	.1504
30	.1666	.1463	.4337	.4319	.4168	.4225

TABLE 22

GROUND COMBAT ATTRITION RATES FOR LTCOLS
CODE 05 ARTILLERY

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	0	0	0	0	0	0
16	.0293	.0310	.0033	.0655	.0508	.0587
17	.1459	.1167	0	.0336	.0362	.0361
18	.1415	.2141	0	.0152	.0192	.0171
19	.1232	.2424	.0136	.0241	.0276	.0262
20	.1475	.2390	.2642	.2813	.2812	.2794
21	.1457	.2282	.0751	.0990	.1051	.1024
22	.1276	.2123	.1129	.1229	.1282	.1257
23	.1361	.1742	.1583	.1996	.2008	.1984
24	.1497	.1446	.3989	.4354	.4264	.4267
25	.1374	.0897	.2138	.2547	.2495	.2472
26	.1187	.0519	.4907	.4585	.4331	.4451
27	.0776	.0448	.8269	.7777	.7420	.7573
28	.0877	.0310	.2152	.2266	.1932	.2102
29	.2426	.1625	.6976	.5877	.5592	.5751
30	.2966	.0207	.6503	.5945	.5703	.5848

TABLE 23

GROUND COMBAT ATTRITION RATES FOR LTCOLS
CODE 10 TANKS

LOS	AGG ORIG	AGG TRANS	MLE	TSCA	JS	LTJS
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	0	0	0	0	0	0
16	.0972	0	.0006	.2593	.2115	.2288
17	.1147	.0077	.0002	.1060	.0935	.0951
18	.1538	.0085	.0001	.0790	.0752	.0746
19	.1315	.0634	0	.0396	.0402	.0404
20	.1032	.0661	.1689	.1690	.1629	.1637
21	.1956	.0548	.0596	.1936	.1939	.1888
22	.1769	.0689	.2050	.2508	.2441	.2422
23	.1716	.0489	.0317	.0899	.0879	.0878
24	.1205	.0429	.1361	.1596	.1460	.1556
25	.1323	.0085	.3439	.4017	.3756	.3854
26	.1995	.0017	.0889	.2828	.2869	.2775
27	.1850	.0077	.1695	.2636	.2373	.2452
28	.1083	.0448	.1714	.4064	.3781	.3828
29	.1252	.0876	.3967	.5706	.5365	.5446
30	.2061	.0004	.6177	.5570	.5312	.5376

V. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

This study investigated the ability of the exponential smoothing model to update attrition rates for Marine Corps manpower models from year to year. Its value as a yearly updating scheme has been demonstrated in this work.

For the reader who wants a single general purpose value of α, we offer the value $\alpha = .4$, but do so reluctantly. Brown has suggested using an α between .01 and .3 in the applications he studied [Ref. 6: p. 106]. Our work, involving manpower attrition rates, appears to call for larger values. However, we have identified some special situations for which the smoothing constant should be considerably larger, e.g., aviation LTCOLS and 1LTS with a pre-1981 base forecasting post-1981 rates.

Exponential smoothing produced estimates for the combat support and ground combat aggregates that were, more often than not, better than those produced by the methods of Major Robinson [Ref. 2], without needing, in theory, the massive data files his methods use to produce the attrition rate estimates. The exponential smoothing model reflected the change in the aviation environment that occurred in 1981 and easily outperformed the Robinson method's estimates in this aggregate for the years following because of the inflexibility of the Robinson method. It also anticipated an unknown source of turbulence in the ground combat first lieutenant attrition rates, and bested the estimates of Robinson's methods for this study block.

Also seen in this study was that three of the four estimators presented by Robinson [Ref. 2], transformed cell scale average, James-Stein, and limited translation James-Stein, outperform the current method used by the

Marine Corps when exponential smoothing is used in all validation years and study blocks except validation years 2 and 3 for combat support 1LTS. These two exceptions, however, were also the only ones seen in Robinson's thesis [Ref. 2: p. 40]. Unfortunately, none of these three emerges as the clear-cut "best" estimator in this study, but all are better than the current Marine Corps estimator.

The maximum likelihood estimator shows the same good performance in transformed scale in this study as it did in Robinson's [Ref. 2: pp. 39-41]. However, there are more cases of large original scale figures of merit for MLE in the present work than were seen previously (see Tables 4-9). These extraordinarily large figures are seen for the ground combat aggregate for both grades and for combat support 1LTS. Not coincidentally, these are the same study blocks for which the use of the optimal α for transformed scale to produce original scale estimates was determined to be unsuitable in the analysis of Appendix C. The ability of the MLE to produce attrition rates better than those of aggregate estimators for the other three study blocks indicates that it, too, may be better than the aggregate methods currently in use, and warrants further study into the cause of the problems in original scale just mentioned. Perhaps if smoothing is done in original scale rather than transformed scale, the optimal properties of MLE will be better displayed.

Thus, Robinson's conclusion [Ref. 2: p. 68] that the MLE, TSCA, JS, and LTJS estimators should be given serious consideration for replacing the Marine Corps' current scheme is reiterated in this work.

B. RECOMMENDATIONS

Despite the encouraging results produced herein, it is not recommended at this time that either exponential smoothing nor any of the four promising estimators presented

as alternatives to the current method be implemented. Further study is needed in the following areas:

1. Base Period. The problem of anticipating major changes in Marine Corps policy such as the one that the aviation OF underwent in 1981 and their effects on the length of the base period should be studied. Also, as more data becomes available, more analyses like the ones conducted herein to determine optimal base period length should be conducted to produce stronger conclusions about base period length.

2. Aviation. As more data is obtained, the aviation aggregate should be re-examined and 1981 used as the first base year to see if lower values of α can be produced, indicating a consistency from year to year in the loss rates which did not exist from 1977 to 1983.

3. Ground Combat 1LTS. The reasons behind the values of α being 1.0, which indicate a year-to-year change in attrition patterns for this study block, should be investigated. Despite the excellent results achieved with respect to the low figures of merit produced, a persistence of such α values would indicate that the base period is totally unimportant, thereby making the use of exponential smoothing unnecessary. The issue is, therefore, whether or not the α values seen for ground combat 1LTS herein are the product of a peculiar set of data. A new analysis of this study block should be undertaken using years other than 1977-1979 as the base period as soon as enough data becomes available to perform validations on newly produced attrition rates.

4. Maximum Likelihood Estimation. In light of the promising performance seen for this estimator in transformed scale using exponential smoothing, an investigation of the unsuitability of using transformed optimal α values to produce original scale loss rates seen for certain study blocks is needed.

5. Aggregation. As recommended by Major Robinson [Ref. 2: p. 69], the work of Amin Elseramegy [Ref. 5], who used the CART routine to find aggregations with encouraging results, should be investigated. Also, the work of Major Tucker [Ref. 1: pp. 75-84], which demonstrated increased attrition at certain lengths of service for certain grades, should be followed up by attempting to aggregate by LOS instead of OF to see if the results can be improved.

APPENDIX A

FUNCTIONS USED IN CALCULATIONS

The following APL functions, and ones similar to them, are used in this study to produce the figures of merit and attrition rates seen in the tables in Chapters III and IV. These procedures require that the following variables be global, i.e., defined throughout the workspace:

(1) N = the estimation period central inventory array, e.g., NA5,

(2) Y = the estimation period loss array, e.g., YA5,

(3) VN = the validation period centarl inventory array, e.g., VNA5,

(4) VY = the validation period loss array, e.g., VYA5,

(5) AN = the estimation period average inventory array, e.g., ANA5,

(6) G = the forced James-Stein shrinkage rate, and

(7) DEE = the factor used in limited translation James-Stein shrinkage.

Functions like A5B seen herein set the values of all of these parameters except G, which is set in the workspace itself to 0.

In order to create the arrays of figures of merit for α values between 0 and 1 for a three-year base period, APL function ALPHAHAT is called (see Figure A.1). ALPHAHAT calls A5B in Figure A.2 to produce the 3-year base estimates from years 1977-1979 and the empirical estimates for 1980. To do this, A5B sets the values of the global variables, then calls ESTIM, seen in Figure A.3, to produce the base estimates. ESTIM, in turn, calls BINPREP in Figure A.4, SUMSQ in Figure A.5, and MLE in Figure A.6 in performing its calculations. A5B then calls XFOUR in Figure A.7 which resets the global variables N, Y, and AN, and calls ESTIM to produce the empirical transformed attrition figures. Notice that the

```
    ∇ ALPHAHAT;LB;UB;STEP;RATES;M1;M2;ALPHA;NUMSTEPS
  ⍝ THIS FUNCTION WILL FIND THE VALUES OF THE RISK
  ⍝ FUNCTION IN BOTH TRANSFORMED AND ORIGINAL SPACE
  ⍝ FOR VARYING VALUES OF ALPHA BETWEEN ANY TWO
  ⍝ LIMITS SET BY THE USER.
  ⍝' INPUT LOWER BOUND'
   LB←0
  ⍝' INPUT UPPER BOUND'
   UB←1
  ⍝' INPUT STEPSIZE'
   STEP←0.02
   ALPHA←LB
   NUMSTEPS←1+(UB-LB)÷STEP
  ⍝ INITIALIZE ARRAYS OR1-OR6 AND TR1-TR6, WHICH
  ⍝ ARE THE ARRAYS OF FOM'S FOR ALPHAS BETWEEN
  ⍝ LB AND UB BY STEPSIZE STEP. VALUES 1-6
  ⍝ CORRESPOND TO AGG ORIG, AGG TRANS, MLE,
  ⍝ TSCA, JS, AND LTJS, RESPECTIVELY.
   OR1←OR2←OR3←OR4←OR5←OR6←(NUMSTEPS,(1↑⍴VN))⍴0
   TR1←TR2←TR3←TR4←TR5←TR6←(NUMSTEPS,(1↑⍴VN))⍴0
   M1←1
  ⍝ CALL A5B TO OBTAIN BASE ESTIMATE S3 AND
  ⍝ EMPIRICAL ESTIMATE S4:
   A5B
  ⍝ HERE IS THE EXPONENTIAL SMOOTHING FUNCTION
  NEXT:R←(ALPHA×X4)+(S3×(1-ALPHA))
  ⍝ CALL RISKO AND RISKT AND PLACE THE RESULTS
  ⍝ FOR THIS ALPHA VALUE IN THE OR AND TR ARRAYS
   RISKO
   RISKT
   M2←1
  ONE:OR1[M1;M2]←R1[M2]
   OR2[M1;M2]←R2[M2]
   OR3[M1;M2]←R3[M2]
   OR4[M1;M2]←R4[M2]
   OR5[M1;M2]←R5[M2]
   OR6[M1;M2]←R6[M2]
   TR1[M1;M2]←RAO[M2]
   TR2[M1;M2]←RAT[M2]
   TR3[M1;M2]←RM[M2]
   TR4[M1;M2]←RT[M2]
   TR5[M1;M2]←RJ[M2]
   TR6[M1;M2]←RL[M2]
   M2←M2+1
   →ONE×⍳(M2≤2)
  NEWSTEP:ALPHA←ALPHA+STEP
   M1←M1+1
   →NEXT×⍳(ALPHA≤UB)
    ∇
```

Figure A.1 APL Function ALPHAHAT

James-Stein shrinkage factors calculated for the base estimates and the DEE values used in limited translation James-Stein for the base estimates are also used for the empirical estimates. ALPHAHAT then smooths the empirical

```
     ∇ A5B
[1]    DEE←0.8
[2]    N← 3 31 1 ↑NA5
[3]    Y← 3 31 1 ↑YA5
[4]    VN←VNA5
[5]    VY←VYA5
[6]    AN← 3 31 1 ↑ANA5
[7]    SS←(⍴VN)⍴1
[8]   ⍝ CALL ESTIM TO PRODUCE BASE ESTIMATE S3
[9]   ⍝ FROM 1977-1979 DATA:
[10]   ESTIM
[11]   S3←R
[12]  ⍝ CALL XFOUR TO PRODUCE THE EMPIRICAL RATES
[13]  ⍝ X4 FOR 1980:
[14]   XFOUR
[15]   G←0
     ∇
```

Figure A.2 APL Function A5B

estimates onto the base estimates and calls RISKO in Figure A.8 and RISKT in Figure A.9 to provide the figures of merit. RISKO calls BINCONV in Figure A.10 to produce original scale loss rates using the inverse arcsine transformation seen in Appendix B for use in the chi-square FOM formula. The resulting arrays, which provide figures of merit for all estimates and for all validation years, can be analyzed to find the α values which minimize transform FOM.

The above example would find the figures of merit for a 3-year base period for aviation 1LTS. In order to obtain the figures of merit for the other five study cells, functions much like A5B are created to set the global variables equal to their respective loss and inventory figures. To make these calculations for 4- and 5-year base period lengths, functions called XFIVE and XSIX are written. These functions are the same as XFOUR except that XFIVE finds the empirical rates for 1981 and XSIX does so for 1982. The A5B-type functions are also changed accordingly to alter the number of years used to calculate the base estimates.

```
∇ P←ESTIM;I;S;U;C1;C;K;M;ZB;ZBBA;AGO;AGT;D
  ⍝ CALCULATES THE AGGREGATE (ORIGINAL),
  ⍝ AGGREGATE (TRANSFORMED), MLE (ORIGINAL),
  ⍝ TSCA (TRANSFORMED), JAMES-STEIN, AND
  ⍝ LIMITED TRANSLATION JAMES-STEIN (LTJS)
  ⍝ ESTIMATORS FOR Z AS SCREENED BY D.
  ⍝ EACH PAGE OF THE R ARRAY IS A SEPARATE
  ⍝ ESTIMATE. THE DEE VALUES USED HEREIN
  ⍝ ARE SET IN THE CALLING PROGRAM A5B,
  ⍝ A8B, ETC.. D IS THE SCREENING
  ⍝ MATRIX CREATED FROM AVERAGE INVENTORY
  ⍝ WHICH SCREENS OUT THE ZERO INVENTORIES.
  ⍝
   Z←Y BINPREP N
   D←(+/ 3 1 2 ⍉AN≠0)≠0
   R←(6,(1↓⍴Z))⍴0
   S←SUMSQ Z
   P←(+/,Y×(⍴Y)⍴D)÷+/,N×(⍴N)⍴D
   AGO←((0.5+(+⌿N÷1↑⍴N))*0.5)×¯1○¯1+2×P
   AGT←(+/,ZB)÷+/,D
   ZBBA←(⍴ZB)⍴ZBB
   G←(G×M≠1)+SHJ×M=1
   C←0⌈1-SHJ←G
   →(2+1↑⎕LC)×⍳0<G
   C←0⌈1-SHJ←(K-3)÷(2-K-K×M)×÷/S
   U←|(ZB-ZBBA)×SHJ*0.5
   R[1;;]←(1↓⍴R)⍴AGO
   R[2;;]←(1↓⍴R)⍴AGT
   R[3;;]←Y MLE N
   R[4;;]←ZB
   R[5;;]←D×ZBBA+C×ZB-ZBBA
 LL:C1←0⌈1-SHL←(1⌊DEE÷U)×SHJ
   R[6;;]←D×ZBBA+C1×ZB-ZBBA
∇
```

Figure A.3 APL Function ESTIM

```
∇ Z←Y BINPREP N
  ⍝ PREPS THE FREEMAN-TUKEY VERSION OF THE
  ⍝ ARC SIN TRANS FOR BINOMIAL DATA
  ⍝ Y IS LOSSES; N IS INVENTORY
   Z←¯1○¯1+2×Y÷N+1
   Z←0.5×((0.5+N)*0.5)×Z+¯1○¯1+2×(Y+1)÷N+1
∇
```

Figure A.4 APL Function BINPREP

Finally, in order to produce the original scale estimates, the A5B-type functions are modified to perform the

```
     ∇ X←SUMSQ Z;SSE
[1]   ⍝ CALCULATES THE SSE AND SSB FOR Z. ALSO
[2]   ⍝ CALCULATES THE MLE (ZB) AND GRAND MEAN
[3]   ⍝ OR AGGREGATE (ZBB), BOTH DIRECTLY FROM
[4]   ⍝ TRANSFORMED DATA.
[5]    K←+/,D
[6]    ZB←D×(+⌿Z)÷M←1↑⍴Z
[7]    ZBB←(+/,ZB)÷K
[8]    X←+/,D×+⌿(Z-(⍴Z)⍴ZBB)*2
[9]    X←(X-SSE),SSE←+/,D×+⌿(Z-(⍴Z)⍴ZB)*2
     ∇
```

Figure A.5 APL Function SUMSQ

```
     ∇ Z←Y MLE N;M1
[1]   ⍝ CALCULATES THE MLE IN THE ORIGINAL SCALE
[2]   ⍝ AND TRANSFORMS IT INTO ARCSIN SPACE.
[3]    D←(+/ 3 1 2 ⍉AN≠0)≠0
[4]    M1←((⍴+⌿Y)⍴D)×(+⌿Y)÷+⌿N
[5]    Z←D×((0.5+(+⌿N)÷1↑⍴N)*0.5)×¯1○¯1+2×M1
     ∇
```

Figure A.6 APL Function MLE

```
     ∇ XFOUR
[1]   ⍝ THIS FUNCTION WILL CALCULATE THE TRANSFORMED
[2]   ⍝ EMPIRICAL ATTRITION RATE X4 FOR USE IN THE
[2]   ⍝ EXPONENTIAL SMOOTHING MODEL FOR ALL SIX
[3]   ⍝ ESTIMATORS. CALLS ESTIM.
[4]    N← ¯1 31 1 ↑NA5
[5]    Y← ¯1 31 1 ↑YA5
[6]    AN← ¯1 31 1 ↑ANA5
[7]    ESTIM
[8]    X4←R
     ∇
```

Figure A.7 APL Function XFOUR

smoothing operation with the α values fixed at their optimal levels for each of the six estimators to produce the

```
     ∇ RISKO;D;K;V;B;AR;S;A;NV
[1]    ⍝ COMPUTES THE ORIGINAL SCALE RISK.
[2]    ⍝ VY AND VN ARE THE VALIDATION
[3]    ⍝ YEAR LOSSES AND INVENTORY. VN, VY,
[4]    ⍝ R, AND AN MUST BE IN THE WORKSPACE.
[5]    ⍝ FIGURES OF MERIT (RISKS) ARE PLACED
[6]    ⍝ IN VECTORS R1 THRU R6, WHICH CORRESPOND
[7]    ⍝ TO ORIG AGG, ORIG TRANS, MLE, TSCA,
[8]    ⍝ JS, AND LTJS, RESPECTIVELY.
[9]    ⍝
[10]     K←1
[11]     D←∨⌿AN>0
[12]     V←((⍴V)⍴D)×V←BINCONV
[13]     AR←((⍴VY)⍴D)×VY÷VN
[14]     S←(V≠0)∧(V≠1)
[15]     R1←R2←R3←R4←R5←R6←(1↑⍴VN)⍴0
[16]   LM:A←(⍴V)⍴AR[K;;]
[17]     NV←(⍴V)⍴VN[K;;]
[18]     RR←((⍴V)⍴D)×S×NV×(S×(A-V)*2)÷V×(1-V)
[19]     R1[K]←+/,SS[K;;]×RR[1;;]
[20]     R2[K]←+/,SS[K;;]×RR[2;;]
[21]     KK1←+/,S[3;;]
[22]     R3[K]←+/,SS[K;;]×RR[3;;]×(KK←+/,D)÷KK1
[23]     R4[K]←+/,SS[K;;]×RR[4;;]
[24]     R5[K]←+/,SS[K;;]×RR[5;;]
[25]     R6[K]←+/,SS[K;;]×RR[6;;]
[26]     K←K+1
[27]     →LM×⍳(K≤1↑⍴VN)
     ∇
```

Figure A.8 APL Function RISKO

smoothed transformed figures. Function BINCONV in Figure A.10 is then called, which inverts the transformation and yields the original scale attrition rate estimates seen in Tables 10 thru 23.

```
     ∇ RISKT;I;D;Z
   ⍝ COMPUTES THE SCALED RISK. VY AND VN ARE THE
   ⍝ VALIDATION YEAR LOSSES AND INVENTORY. VN, VY,
   ⍝ R, AND AC MUST BE IN THE WORKSPACE.
   ⍝ FIGURES OF MERIT (RISKS) ARE PLACED
   ⍝ IN VECTORS RAO, RAT, RM, RT, RJ, AND
   ⍝ RL, WHICH CORRESPOND TO
   ⍝ ORIG AGG, ORIG TRANS, MLE, TSCA,
   ⍝ JS, AND LTJS, RESPECTIVELY.
   ⍝
    Z←VY BINPREP VN
    I←1
    RAO←RAT←RM←RT←RJ←RL←(1↑ρVN)ρ0
    D←(+/ 3 1 2 ⍉AN≠0)≠0
  LL:RAO[I]←(+/,SS[I;;]×(Z[I;;]-R[1;;])*2)÷+/,D
    RAT[I]←(+/,SS[I;;]×(Z[I;;]-R[2;;])*2)÷+/,D
    RM[I]←(+/,SS[I;;]×(Z[I;;]-R[3;;])*2)÷+/,D
    RT[I]←(+/,SS[I;;]×(Z[I;;]-R[4;;])*2)÷+/,D
    RJ[I]←(+/,SS[I;;]×(Z[I;;]-R[5;;])*2)÷+/,D
    RL[I]←(+/,SS[I;;]×(Z[I;;]-R[6;;])*2)÷+/,D
    I←I+1
    →LL×⍳(I≤(1↑ρVN))
     ∇
```

Figure A.9 APL Function RISKT

```
     ∇ BINCONV;V0;V1;N1;D
   ⍝ INVERTS ARC SIN TRANSFORMATION. R IS THE
   ⍝ ARRAY TO BE TRANSFORMED. N IS THE CENTRAL
   ⍝ INVENTORY FOR THE ESTIMATION YEARS.
    D←∨⌿AN>0
    B←0.5×1+1○R÷(0.5+(N1←(ρR)ρ+⌿N)÷1↑ρN)*0.5
    V0←R<(-○÷2)×(N1+0.5)*0.5
    V1←R>(○÷2)×(N1+0.5)*0.5
    B←((ρR)ρD)×V1+B×V0∨V1
   ⍝⍉ 6 31 ρ B
     ∇
```

Figure A.10 APL Function BINCONV

APPENDIX B

FREEMAN-TUKEY ARCSINE TRANSFORMATION

1. GENERAL

Because the TSCA, James-Stein and limited translation James-Stein techniques make the assumptions that the distribution of the number of losses is normally distributed with constant variance, and because the binomial model for the loss data does not meet these assumptions, a transformation is needed. The Freeman-Tukey arcsine transformation produces values for which normality and constant variance become more tenable assumptions.

Robinson demonstrated in his thesis [Ref. 2: pp. 74-79] that both the normality and variance assumptions are compromised somewhat for low values of n and p. He therefore concluded that the Freeman-Tukey transform is unreliable at such values. Therefore, the validity of the results for James-Stein estimation and limited translation James-Stein must be questioned in this analysis as they were in Robinson's [Ref. 2: p. 19]. The following two equations are represented in APL by functions BINPREP (transform) and BINCONV (inverse transform). See Appendix A.

2. THE TRANSFORMATION

The equation for the transformation is:

$$x = 0.5(n+0.5)^{1/2} \sin^{-1}(2y/(n+1)-1) + \sin^{-1}(2(y+1)/(n+1)-1) \qquad (B.1)$$

This equation transforms raw losses, y, into transformed losses, x using the central inventory, n.

3. THE INVERSE TRANSFORMATION

To invert the transformation and produce the rates in original space, use the following set of equations:

$$n_{ij} = (1/T)\sum n_{ij}(t), \text{ for all } i \tag{B.2}$$

$$v_{ij} = x_{ij}/(n_{ij}+.5).5 \tag{B.3}$$

$$r = \begin{cases} 0 & v_{ij} \text{ .le. } -\pi/2 \\ .5(1+\sin v_{ij}(t)) \quad \text{if} & -\pi/2<v_{ij}<\pi/2 \\ 1 & v_{ij} \text{ .ge. } \pi/2 \end{cases} \tag{B.4}$$

where n_{ij} is the central inventory for the i^{th} LOS and the j^{th} OF, x_{ij} is the corresponding transformed attrition figure, and v_{ij} the corresponding loss rate estimates in the original scale.

APPENDIX C

ANALYSIS OF OPTIMAL ALPHA OF TRANSFORMED AND ORIGINAL SCALES

1. GENERAL

The following tables give the optimal values of α for transformed and original scales for the three-year base period used in the production of the attrition rate estimates in Chapter IV. The values of α listed are those which produce the minimum figures of merit for validation year 2, 1982.

As one can see from the comparison tables, the α values producing the minimum figures of merit for transformed scale are very close in most cases to the α values producing the minimum figures of merit for original scale, with the notable exceptions being MLE for combat support and ground combat. Each of the aggregates is discussed below.

TABLE 24

COMPARISON OF TRANSFORMED AND ORIGINAL ALPHA
AVIATION AGGREGATE

1LTS	TRANSFORMED α	ORIGINAL α
AGG ORIG	1.00	1.00
AGG TRANS	.00	1.00
MLE	1.00	1.00
TSCA	1.00	1.00
JS	1.00	.98
LTJS	1.00	.98
LTCOLS	**TRANSFORMED α**	**ORIGINAL α**
AGG ORIG	1.00	1.00
AGG TRANS	1.00	1.00
MLE	1.00	.98
TSCA	1.00	1.00
JS	1.00	1.00
LTJS	1.00	1.00

Note: Optimal α values are for validation year 1982.

2. AVIATION

Looking at Table 24, one sees that for aviation, the α values match up almost perfectly except for the transformed aggregate for 1LTS, which has values at completely opposite ends of the limits of α for the two scales. However, since the aggregate transform estimator shows very poor performance throughout the analyses in Chapters III and IV, with much larger figures of merit than the other estimators, inconsistencies like this are not of much concern.

TABLE 25

COMPARISON OF TRANSFORMED AND ORIGINAL ALPHA
COMBAT SUPPORT AGGREGATE

1LTS	TRANSFORMED α	ORIGINAL α
AGG ORIG	1.00	1.00
AGG TRANS	.00	.00
MLE	.56	1.00
TSCA	.60	.50
JS	.90	.90
LTJS	.64	.56
LTCOLS	**TRANSFORMED α**	**ORIGINAL α**
AGG ORIG	.00	.00
AGG TRANS	1.00	1.00
MLE	.38	.34
TSCA	.34	.36
JS	.26	.36
LTJS	.30	.36

Note: Optimal α values are for validation year 1982.

3. COMBAT SUPPORT

From an analysis of Table 25, one sees that the combat support aggregate likewise shows a consistency in the values of α for transformed and original scales. The only exceptions to this are MLE for 1LTS, which has a difference in α values of a whopping .44, and, to a much lesser extent, TSCA for 1LTS and JS for LTCOLS, which each have a difference in α values of .10.

TABLE 26

COMPARISON OF TRANSFORMED AND ORIGINAL ALPHA
GROUND COMBAT AGGREGATE

1LTS	TRANSFORMED α	ORIGINAL α
AGG ORIG	.92	1.00
AGG TRANS	.00	.34
MLE	.94	.64
TSCA	1.00	.92
JS	1.00	1.00
LTJS	1.00	.94
LTCOLS	**TRANSFORMED α**	**ORIGINAL α**
AGG ORIG	1.00	1.00
AGG TRANS	.00	.00
MLE	.42	1.00
TSCA	.56	.52
JS	.58	.50
LTJS	.56	.50

Note: Optimal α values are for validation year 1982.

4. GROUND COMBAT

The ground combat aggregate also shows consistency between optimal values of α in transformed and original scales for all estimators except MLE for both grades and AGG TRANS for 1LTS. The lack of consistency between the α values for transformed and original scales for MLE seen in Table 26 as well as in the 1LTS section of Table 25 is of major concern. Unfortunately, using the transformed optimal α value to produce figures of merit in the original scale for MLE has a big effect on those figures, making them much larger. This casts into doubt the ability of the exponential smoothing model to produce good maximum likelihood estimates of attrition rates.

LIST OF REFERENCES

1. Tucker, D.D., *Loss Rate Estimation in Marine Corps Officer Manpower Models*, Masters Thesis, Naval Postgraduate School, Monterey, California, September 1985.

2. Robinson, John R., *Limited Translation Shrinkage Estimation of Loss Rates in Marine Corps Manpower Models*, Masters Thesis, Naval Postgraduate School, Monterey, California, March 1986.

3. Efron, B. and Morris, C., "Limiting the Risk of Bayes and Empirical Bayes Estimators -- Part I: The Bayes Case," Journal of the American Statistical Association, v. 66, pp. 807-815, December 1971.

4. Efron, B. and Morris, C., "Limiting the Risk of Bayes and Empirical Bayes Estimators -- Part II: The Empirical Bayes Case," Journal of the American Statistical Association, v. 67, pp. 130-139, March 1972.

5. Amin Elseramegy, H., *CART Program: Implementation of the CART Program and Its Application to Estimating Attrition Rates*, Masters Thesis, Naval Postgraduate School, Monterey, California, December 1985.

6. Brown, Robert Goodell, *Smoothing, Forecasting and Prediction of Discrete Time Series*, Englewood Cliffs, N.J. : Prentice-Hall, 1963.

7. Bartholomew, David J. and Forbes, Andrew F., *Statistical Techniques for Manpower Planning*, Norwich, Great Britain: Wiley, 1979.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22304-6145	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93943-5000	2
3.	Professor Robert R. Read, Code 55Re Naval Postgraduate School Monterey, California 93943-5000	5
4.	Lieutenant Colonel Jack B. Gafford, Code 55Gf Naval Postgraduate School Monterey, California 93943-5000	1
5.	Professor Paul R. Milch, Code 55Mh Naval Postgraduate School Monterey, California 93943-5000	1
6.	Marine Corps Representative, Code 0309 Naval Postgraduate School Monterey, California 93943-5000	1
7.	Commandant of the Marine Corps HQMC, Code MPP-30 Washington, D.C. 22314	1
8.	Commandant of the Marine Corps HQMC, Code MPP-32 Washington, D.C. 22314	1
9.	Commandant of the Marine Corps HQMC, Code MPI-10 Washington, D.C. 22314	1
10.	Commandant of the Marine Corps HQMC, Code MPI-20 Washington, D.C. 22314	1
11.	Commandant of the Marine Corps HQMC, Code MPI-40 Washington, D.C. 22314	1
12.	Commanding Officer Navy Personnel Research and Development Center San Diego, California 92152	1
13.	First Lieutenant Daniel L. Hogan, Jr. 36 Meade Avenue Hicksville, New York 11801	5